Huangxin Chen, Xuejun Xu, Ronald H.W. Hoppe

Convergence and Optimality of Adaptive Nonconforming Finite Element Methods for Nonsymmetric and Indefinite Problems

Preprint Nr. 11/2008 — 28. Februar 2008
Institut für Mathematik, Universitätsstraße, D-86 135 Augsburg http://www.math.uni-augsburg.de/

Impressum:

Herausgeber:
Institut für Mathematik
Universität Augsburg
86135 Augsburg
`http://www.math.uni-augsburg.de/forschung/preprint/`

ViSdP:
Ronald H.W. Hoppe
Institut für Mathematik
Universität Augsburg
86135 Augsburg

CONVERGENCE AND OPTIMALITY OF ADAPTIVE NONCONFORMING FINITE ELEMENT METHODS FOR NONSYMMETRIC AND INDEFINITE PROBLEMS

HUANGXIN CHEN* , XUEJUN XU*, AND R.H.W. HOPPE†

Abstract. Recently an adaptive nonconforming finite element method (ANFEM) has been developed by Carstensen and Hoppe in [9]. In this paper, we extend their result to nonsymmetric and indefinite problems. The main tools in our analysis are a posteriori error estimators and a quasi-orthogonality property. In this case, we need to overcome two main difficulties: one stems from the nonconformity of the finite element space, the other is how to handle the effect of a nonsymmetric and indefinite bilinear form. Two ANFEM algorithms (ANFEM I, ANFEM II) are proposed for the lowest order Crouzeix-Raviart element. It is shown that both ANFEM algorithms are a contraction for the sum of the energy error and a scaled volume term between two consecutive adaptive loops. Moreover, optimality in the sense of optimal algorithmic complexity can be shown for ANFEM II. The results of numerical experiments confirm the theoretical findings.

1. Introduction.

A standard adaptive finite element method (AFEM) consists of successive loops of the cycle

$$\textbf{SOLVE} \longrightarrow \textbf{ESTIMATE} \longrightarrow \textbf{MARK} \longrightarrow \textbf{REFINE}. \tag{1.1}$$

A posteriori error estimation is crucial in the AFEM procedure. This subject has been extensively studied for conforming finite element methods. As far as nonconforming finite element methods are concerned, Dari, Duran, Padra and Vampa [21] presented an a posteriori error estimator for a model problem based on a residual estimator supplemented by tangential components of the jumps of the fluxes across interior element edges. They showed that the estimator, together with a volume term, provides both an upper and a lower bound for the energy error. Hierarchical-type estimators for nonconforming finite element approximations were derived and analyzed by Hoppe and Wohlmuth [25]. In the same spirit, Schieweck [32] established an a posteriori error estimator featuring an additive post-processing term as in [25]. Using averaging techniques, Carstensen, Bartels and Jansche suggested Zienkiewicz/Zhu-type a posteriori error estimators for nonconforming finite elements in [13]. Moreover, Ainsworth [1] extended the equilibrated residual method to the a posteriori error analysis of nonconforming finite element approximations of linear second order elliptic equations with diffusion coefficients undergoing large jumps across interfaces. Recently, Carstensen, Hu and Orlando [11] provided a unified framework for the a posteriori error analysis of a large class of nonconforming finite element methods (cf. also Carstensen and Hu [12]).

Besides convergence, optimality is another important issue in AFEM which was first addressed by Binev, Dahmen, DeVore [7] and further studied by Stevenson [35], who showed optimality without additional coarsening which was required in [7]. Very recently, Cascon, Kreuzer, Nochetto, Siebert [15] succeeded to establish optimality of AFEM without the assumption of an interior node property. Other extensions have been considered by Carstensen and Hoppe in [10] and Chen, Holst and Xu in [17] for mixed AFEM, and by Becker, Mao and Shi for a simple ANFEM in [5].

*LSEC, Institute of Computational Mathematics, Chinese Academy of Sciences, P.O.Box 2719, Beijing, 100080, People's Republic of China(chx@lsec.cc.ac.cn xxj@lsec.cc.ac.cn).

†Institut für Mathematik, Universität Augsburg, D-86159, Augsburg,Germany, and Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA(rohop@math.uh.edu).

Given a bounded, polyhedral domain $\Omega \subset R^d$, $d = 2$ or 3, we consider the following nonsymmetric and indefinite elliptic boundary value problem

$$\mathcal{L}u := -\mathrm{div}(\mathbf{A}\nabla u) + \mathbf{b}\cdot\nabla u + cu = f \qquad \text{in}\,\Omega, \tag{1.2}$$

$$u = 0 \qquad \text{on}\partial\Omega. \tag{1.3}$$

The choice of a homogeneous Dirichlet boundary condition is made for ease of presentation only. Similar results are valid for other types of boundary conditions as well. We further assume that the coefficient functions in (1.2) satisfy the following properties:

(**a**) $\mathbf{A}$: $\Omega \longrightarrow R^{d\times d}$ is a Lipschitz and symmetric positive definite matrix, and there exist $0 < \mu < M < \infty$ such that for all $\xi \in R^d$ and almost all $x \in \Omega$ there holds

$$\mu|\xi|^2 \le \mathbf{A}(x)\xi\cdot\xi \le M|\xi|^2, \tag{1.4}$$

where $|\cdot|$ stands for the Euclidean norm;
(**b**) $\mathbf{b} \in [L^\infty(\Omega)]^d$;
(**c**) $c \in L^\infty(\Omega)$ is nonnegative, i.e., $c \ge 0$ a.e. in Ω;
(**d**) $f \in L^2(\Omega)$.

The objective of this paper is to study the convergence and optimality for a sequence of discrete solutions computed by adaptive nonconforming finite element methods (ANFEM) of the cycle (1.1) applied to the above nonsymmetric and indefinite problem (1.2),(1.3). We note that recently Mekchay and Nochetto [28] proposed an AFEM for these problems based on conforming finite element methods and proved convergence. Nonconforming finite element methods are known to have better stability properties for certain classes of problems, e.g., convection-dominated convection-diffusion problems. Another motivation to consider nonconforming methods for general elliptic problems is that they exhibit a close relationship to mixed methods. This relationship could be further exploited to derive efficient numerical algorithms for the mixed formulation as well. There exist several contributions regarding nonconforming and mixed finite elements for nonsymmetric and indefinite problems (cf. [3, 16, 19, 22, 31, 34, 37, 38, 39, 40]). In this paper, we shall extend the ANFEM developed by Carstensen and Hoppe in [9] to nonsymmetric and indefinite elliptic boundary value problems. The main tools in our a posteriori error analysis are reliability and efficiency along with a quasi-orthogonality property. To this end, we need to overcome two main difficulties: one stems from the nonconformity of the approach, the other is due to the impact of the nonsymmetric and indefinite character of the problem.

The main results of this paper are the following convergence and optimality results for two ANFEM algorithms, referred to as ANFEM I and ANFEM II: Let u_h^N,u_H^N be two successive finite element solutions obtained by the ANFEM algorithms involving loops of the cycle (1.1). Under the assumptions that the matrix-valued function $\mathbf{A}$ is piecewise constant, and the initial mesh size h_0 is sufficiently small, there exist positive constants $0 < \rho < 1$ and $\beta > 0$, depending on the shape regularity of meshes, the coefficients of the equation (1.2), and the parameters used by ANFEM such that

$$|||u - u_h^N|||_h^2 + \beta J_h^2 \le \rho(|||u - u_H^N|||_H^2 + \beta J_H^2).$$

Here, $|||\cdot|||_h$ stands for the energy norm. The volume terms J_h, J_H induced by the operator $\mathcal{L}$ will be defined later. We emphasize that in the proof of the reliability and the efficiency of the a posteriori error estimators we do not need to require the matrix-valued function $\mathbf{A}$ to be piecewise constant. However, the quasi-orthogonality result will heavily rely on this assumption. In addition, the initial mesh is chosen to reflect the structure of $\mathbf{A}$, i.e., $\mathbf{A}$ is piecewise constant on $\mathcal{T}_0$. We also assume that $\mathbf{A}$ does not undergo large jumps in Ω. The situation where $\mathbf{A}$ exhibits large jumps in Ω will be addressed in the last section.

As far the complexity of ANFEM is concerned, we obtain the following result for the algorithm ANFEM II featuring a marking strategy based on the comparison of the a posteriori error estimator and a volume term:

$$(|||u - u_h^N|||_h^2 + \beta J_h^2)^{\frac{1}{2}} \lesssim (N_h - N_0)^{-s}.$$

Here, N_h refers to the number of degrees of freedom of the nonconforming finite element space V_h^N. Assuming (u, f, D) belongs to some approximation class $\mathbb{A}_s$ where the index s is used to characterize the best possible approximation rate of u in the norm $|||\cdot|||_h$. s depends on the regularity of the solution, the data and the polynomial order of the finite elements.

The remaining part of this paper is organized as follows: In section 2, we introduce some notations, recall the existence and uniqueness of solutions of the variational formulation of (1.2), (1.3) and its associated nonconforming approximation. Section 3 discusses the reliability and efficiency of a posteriori error estimators, whereas section 4 is devoted to the quasi-orthogonality property. We propose the algorithm ANFEM I and prove its convergence in section 5. The optimal complexity of the algorithm ANFEM II is shown in section 6. In section 7, we present some numerical experiments to illustrate the performance of ANFEM I and ANFEM II. Finally, in the last section we extend the results to the case of coefficient functions with large jumps in the domain.

2. Notations and Preliminaries. Throughout this paper, we adopt standard notations from Lebesgue and Sobolev space theory (cf., e.g., [20]). In particular, we refer to $(\cdot,\cdot)_{0,\Omega}$, $\|\cdot\|_{0,\Omega}$ as the inner product and norm on $L^2(\Omega)$ and to $\|\cdot\|_{1,\Omega}$ as the norm on the Sobolev space $H^1(\Omega)$. We further use $A \lesssim B$, if $A \leq CB$ with a positive constant C depending only on the shape regularity of the meshes, the coefficients in (1.2), and the parameters used by ANFEM. $A \approx B$ stands for $A \lesssim B \lesssim A$. We restrict ourselves to the 2D case. The analysis of 3D problems is similar.

The weak formulation of (1.2) and (1.3) amounts to the computation of $u \in V := H_0^1(\Omega)$ such that

$$\mathcal{B}(u, v) = (f, v) \quad , \quad v \in V, \tag{2.1}$$

where the bilinear form $\mathcal{B} : V \times V \to \mathbb{R}$ is given by

$$\mathcal{B}(u, v) = (\mathbf{A}\nabla u, \nabla v) + (\mathbf{b} \cdot \nabla u + cu, v) \quad , \quad u, v \in V. \tag{2.2}$$

We denote by $|||\cdot|||$ the energy norm

$$|||v|||^2 = (\mathbf{A}\nabla v, v) + (cv, v) \quad , \quad v \in V,$$

which is equivalent to the $||\cdot||_{1,\Omega}$-norm.

In case the bilinear form (2.2) is bounded and V-elliptic, the existence and uniqueness of a solution of (2.1) follows from the Lemma of Lax-Milgram. If only Gårding's inequality

$$\mathcal{B}(v,v) + \lambda_1 ||v||_{0,\Omega}^2 \geq |||v|||^2 \quad , \quad v \in V \tag{2.3}$$

holds true with λ_1 depending on the coefficients $\mathbf{A}$, $\mathbf{b}$, c, we refer to [23] and [24] for an existence and uniqueness result.

We further suppose that the domain Ω and the data are such that the weak solution of (2.1) satisfies

$$u \in V \cap H^{1+\alpha}(\Omega) \quad \text{for some } 0 < \alpha \leq 1. \tag{2.4}$$

Throughout this paper, we work with families of shape regular meshes $\{\mathcal{T}_h\}$. The set of edges is denoted by $\mathcal{E}_h$, and the set of interior and boundary edges by $\mathcal{E}_h^0$ and $\mathcal{E}_h^{\partial\Omega}$, respectively. Correspondingly, we refer to $\mathcal{N}_h$ as the set of nodes and to $\mathcal{N}_h^0$ as the set of interior nodes. For any $E \in \mathcal{E}_h$, h_E and m_E denote the length and the midpoint of E. The patch $\omega_E, E \in \mathcal{E}_h^0$, is the union of two elements in $\mathcal{T}_h$ sharing E. For any $T \in \mathcal{T}$, h_T and x_T stand for the diameter and the barycenter of T. The domain Ω_z is the union of elements containing $z \in \mathcal{N}_h$ and h_z refers to the diameter of Ω_z.

We denote by V_h^N the lowest order nonconforming Crouzeix-Raviart finite element space with respect to $\mathcal{T}_h$, i.e.,

$$V_h^N = \{v_h^N \in L^2(\Omega) \mid v_h^N|_T \in P_1(T), T \in \mathcal{T}_h \ , \ \int_E [v_h^N] ds = 0, E \in \mathcal{E}_h\}.$$

Here, $[v_h^N]|_E$ refers to the jump of v_h^N across $E \in \mathcal{E}_h^0$ and is set to zero for $E \in \mathcal{E}_h^{\partial\Omega}$. Further, we define the conforming P_1 finite element space by

$$V_h^c = \{v_h^c \in V \mid v_h|_T \in P_1(T), T \in \mathcal{T}_h\}.$$

The nonconforming finite element approximation of (2.1) is to find $u_h^N \in V_h^N$ such that

$$\mathcal{B}_h(u_h^N, v_h^N) = (f, v_h^N) \quad , \quad v_h^N \in V_h^N, \tag{2.5}$$

where $\mathcal{B}_h(\cdot,\cdot)$ stands for the discrete bilinear form

$$\begin{aligned} \mathcal{B}_h(u_h^N, v_h^N) &= \sum_{T \in \mathcal{T}_h} ((\mathbf{A}\nabla u_h^N, \nabla v_h^N)_{0,T} + (\mathbf{b}\cdot\nabla u_h^N + cu_h^N, v_h^N)_{0,T}) \\ &:= (\mathbf{A}\nabla_h u_h^N, \nabla_h v_h^N) + (\mathbf{b}\cdot\nabla_h u_h^N + cu_h^N, v_h^N). \end{aligned} \tag{2.6}$$

Here, ∇_h denotes the elementwise gradient with respect to $\mathcal{T}_h$. Likewise, we refer to $|||\cdot|||_h$ as the mesh-dependent energy norm

$$|||v_h^N|||_h^2 := (\mathbf{A}\nabla_h v_h^N, \nabla_h v_h^N) + (cv_h^N, v_h^N). \tag{2.7}$$

For sufficiently small h, (2.5) admits a unique solution (cf., e.g., [16]). We set $W = V \oplus V_h^N$ and suppose

$$\mathcal{B}_h(w,w) + \lambda ||w||_{0,\Omega}^2 \geq |||w|||_h^2 \quad , \quad w \in W, \tag{2.8}$$

which due to (**a-d**) can be assumed to hold true.

3. Reliability and efficiency of a posteriori error estimators.

3.1. Reliability of a posteriori error estimators. We focus on the reliability of an a posteriori error estimator which allows control of the error due to the approximation of the solution of (2.1) by a discrete solution of (2.5). We note that the reliability of a posteriori error estimators for nonconforming finite element approximations has been studied in [1, 2, 8, 9, 11, 12, 13, 21]. Here, we will extend these results to the nonsymmetric and indefinite problem (1.2),(1.3).

The following weighted Clément-type interpolation operator suggested in [8] is crucial for the proof of the reliability.

DEFINITION 3.1. ([8]) *Define a linear mapping* $\mathcal{J} : L^1(\Omega) \longrightarrow V_h^c$ *via*

$$\mathcal{J}f := \sum_{z\in N_h^0} \frac{(f,\psi_z)}{(1,\varphi_z)} \cdot \varphi_z \quad , \quad f \in L^1(\Omega) \quad , \quad \psi_z := \frac{\varphi_z}{\sum_{z\in N_h^0} \varphi_z},$$

where $\{\varphi_z | z \in N_h\}$ *is a Lipschitz partition of unity on* Ω.

LEMMA 3.1. ([8]) *For the operator* $\mathcal{J} : V \longrightarrow V_h^c$ *there holds*

$$||\nabla \mathcal{J}\phi||_{0,\Omega} + \sum_{T\in\mathcal{T}_h} h_T^{-1} ||\phi - \mathcal{J}\phi||_{0,T} + \sum_{E\in\mathcal{E}_h^0} h_E^{-\frac{1}{2}} ||\phi - \mathcal{J}\phi||_{0,E} \lesssim ||\nabla\phi||_{0,\Omega} \quad , \quad \phi \in V.$$

Moreover, for all $f \in L^2(\Omega)$

$$\int_\Omega f(\phi - \mathcal{J}\phi)dx \lesssim ||\nabla\phi||_{0,\Omega} (\sum_{z\in N_h^0} h_z^2 ||f - \overline{f_z}||_{0,\Omega_z})^{\frac{1}{2}} \quad , \quad \phi \in V,$$

where $\overline{f_z}$ *stands for the mean value of* f *with respect to* Ω_z.

The step ESTIMATE of the adaptive loop (1.1) consists of computing an a posteriori error estimator. For $E \in \mathcal{E}_h$ we define the edge residual of the tangential component according to

$$\eta_{s,E}(v_h^N) := \begin{cases} h_E^{\frac{1}{2}} ||[\frac{\partial v_h^N}{\partial s}]||_{0,E} \ , \ E \in \mathcal{E}_h^0 \\ h_E^{\frac{1}{2}} ||\frac{\partial v_h^N}{\partial s}||_{0,E} \ , \ E \in \mathcal{E}_h^{\partial\Omega} \end{cases} , \tag{3.1}$$

and we set

$$\eta_s^2(u_h^N, \mathcal{F}) := \sum_{E\in\mathcal{F}} \eta_{s,E}^2(u_h^N) \quad , \quad \mathcal{F} \subseteq \mathcal{E}_h. \tag{3.2}$$

Likewise, the edge residual of the normal component is defined by means of

$$\eta_{\nu,E}(u_h^N) := h_E^{\frac{1}{2}} ||[\mathbf{A}\nabla u_h^N] \cdot \nu_E||_{0,E} \quad , \quad E \in \mathcal{E}_h^0, \tag{3.3}$$

where ν refers to the unit outward normal vector on E, and we set

$$\eta_\nu^2(u_h^N, \mathcal{F}) := \sum_{E\in\mathcal{F}} \eta_{\nu,E}^2(u_h^N) \quad , \quad \mathcal{F} \subseteq \mathcal{E}_h^0. \tag{3.4}$$

Moreover, we consider the element residuals

$$R_T(u_h^N) := f + \operatorname{div}(\mathbf{A}\nabla u_h^N) - \mathbf{b}\cdot\nabla u_h^N - cu_h^N \quad , \quad T \in \mathcal{T}_h.$$

The volume term and the oscillation term are given by

$$J_T(u_h^N) := h_T||R_T(u_h^N)||_{0,T} \quad , \quad T \in \mathcal{T}_h, \tag{3.5}$$

$$\mathrm{osc}_T(u_h^N) := h_T||R_T(u_h^N) - \overline{R_T(u_h^N)}||_{0,T} \quad , \quad T \in \mathcal{T}_h, \tag{3.6}$$

where $\overline{R_T(u_h^N)}$ is the mean value of $R_T(u_h^N)$ with respect to T. We set

$$J_h^2(u_h^N, \mathcal{M}) := \sum_{T\in\mathcal{M}} J_T^2(u_h^N) \quad , \quad \mathcal{M} \subseteq \mathcal{T}_h,$$

$$\mathrm{osc}_h^2(u_h^N, \mathcal{M}) := \sum_{T\in\mathcal{M}} \mathrm{osc}_T^2(u_h^N) \quad , \quad \mathcal{M} \subseteq \mathcal{T}_h.$$

For ease of notation we further set

$$\begin{aligned} \eta_s^2(u_h^N) &:= \eta_s^2(u_h^N, \mathcal{E}_h) \quad , \quad \eta_\nu^2(u_h^N) := \eta_\nu^2(u_h^N, \mathcal{E}_h^0), \\ J_T &:= J_T(u_h^N) \qquad , \quad J_h^2(\mathcal{M}) := J_h^2(u_h^N, \mathcal{M}), \\ \mathrm{osc}_T &:= \mathrm{osc}_T(u_h^N) \quad , \quad \mathrm{osc}_h^2(\mathcal{M}) := \mathrm{osc}_h^2(u_h^N, \mathcal{M}), \end{aligned}$$

and

$$J_h^2 := J_h^2(\mathcal{T}_h) \quad , \quad \mathrm{osc}_h^2 := \mathrm{osc}_h^2(\mathcal{T}_h).$$

The following two lemmas will be used to prove the reliability of the estimator.

LEMMA 3.2 (Helmholtz Decomposition [13]). *There exist $\phi \in V$ and $\psi \in H^1(\Omega)$ with $\mathrm{curl}\psi \cdot \nu|_{\partial\Omega} = 0$ such that*

$$\boldsymbol{A}\nabla_h(u - u_h^N) = \boldsymbol{A}\nabla\phi + \mathrm{curl}\psi. \tag{3.7}$$

LEMMA 3.3. *For all $w \in V$ there holds*

$$\min_{w\in V} ||\boldsymbol{A}^{\frac{1}{2}}\nabla_h(w - u_h^N)||_{0,h}^2 \lesssim \eta_s^2(u_h^N). \tag{3.8}$$

Proof. Applying the Helmholtz decomposition from Theorem 3.1 in [12], we can easily prove that for a Lipschitz continuous partition of unity $\{\psi_z, z \in N_h^0\}$,

$$||\mathbf{A}^{\frac{1}{2}}\nabla_h(w - u_h^N)||_{0,h}^2 \lesssim \sum_{z\in N_h^0} \sum_{E\in\mathcal{E}_h(\Omega_z)} h_E||[\frac{\partial(\psi_z u_h^N)}{\partial s}]||_{0,E}^2 \quad , \quad w \in V. \tag{3.9}$$

Moreover, we have

$$[\frac{\partial(\psi_z u_h^N)}{\partial s}]|_E = [\nabla(\psi_z u_h^N)] \cdot \tau_E = \nabla\psi_z \cdot \tau_E[u_h^N] + \psi_z[\nabla_h u_h^N] \cdot \tau_E,$$

and

$$||[u_h^N]||_{0,E}^2 \approx h_E([u_h^N]^2(m_1) + [u_h^N]^2(m_2)),$$

where m_1 and m_2 are the Gaussian quadrature points on E. Observing $[u_h^N](m_E) = 0$, we obtain

$$u_h^N|_{T_1}(a_E) - u_h^N|_{T_2}(a_E) = \frac{h_E}{2}[\frac{\partial u_h^N}{\partial s}], \tag{3.10}$$

where a_E is one of the end points of the edge E. Furthermore, observing that $u_h^N(m_i), 1 \le i \le 2$, is a linear combination of $u_h^N(a_E)$ and $u_h^N(m_E)$, we deduce

$$||[u_h^N]||_{0,E} \lesssim h_E||[\frac{\partial u_h^N}{\partial s}]||_{0,E}.$$

Combining the previous inequalities and using $||\nabla \psi_z||_{L^\infty(E)} \lesssim h_E^{-1}$, it follows that

$$||[\frac{\partial(\psi_z u_h^N)}{\partial s}]||_{0,E} \lesssim ||[\frac{\partial u_h^N}{\partial s}]||_{0,E},$$

which, together with (3.9), yields (3.8). ☐

THEOREM 3.1 (Reliability). *Under the assumption that the initial mesh size h_0 is sufficiently small there holds*

$$|||u - u_h^N|||_h \lesssim \mathrm{osc}_h + \eta_s(u_h^N) + \eta_\nu(u_h^N). \tag{3.11}$$

Proof. In view of (2.6) we have

$$\begin{aligned}\mathcal{B}_h(u - u_h^N, u - u_h^N) = {} & (\mathbf{A}\nabla_h(u - u_h^N), \nabla_h(u - u_h^N)) \\ & +(\mathbf{b}\cdot\nabla_h(u - u_h^N) + c(u - u_h^N), u - u_h^N).\end{aligned}$$

Lemma 3.2 tells us

$$\mathbf{A}\nabla_h(u - u_h^N) = \mathbf{A}\nabla\phi + \mathrm{curl}\psi \quad , \quad \phi \in V, \psi \in H^1(\Omega),$$

and hence,

$$\begin{aligned}\mathcal{B}_h(u - u_h^N, u - u_h^N) = {} & (\mathbf{A}\nabla\phi + \mathrm{curl}\psi, \nabla_h(u - u_h^N)) \\ & +(\mathbf{b}\cdot\nabla_h(u - u_h^N) + c(u - u_h^N), u - u_h^N).\end{aligned}$$

Using the weighted Clément-type interpolation operator from Definition 3.1, we obtain

$$\begin{aligned}&(\mathbf{A}\nabla\phi, \nabla_h(u - u_h^N)) = (f + \mathrm{div}(\mathbf{A}\nabla_h u_h^N), \phi - \mathcal{J}\phi) \\ &+ \sum_{E\in\mathcal{E}_h^0} \int_E [\mathbf{A}\nabla_h u_h^N \cdot \nu_E](\phi - \mathcal{J}\phi)ds + (\mathbf{b}\cdot\nabla_h u_h^N + cu_h^N, \mathcal{J}\phi) - (\mathbf{b}\cdot\nabla u + cu, \phi).\end{aligned}$$

Observing

$$\begin{aligned}&(\mathbf{b}\cdot\nabla_h(u - u_h^N) + c(u - u_h^N), u - u_h^N) + (\mathbf{b}\cdot\nabla_h u_h^N + cu_h^N, \mathcal{J}\phi) - (\mathbf{b}\cdot\nabla u + cu, \phi) \\ &= (\mathbf{b}\cdot\nabla_h(u - u_h^N) + c(u - u_h^N), u - u_h^N - \phi) - (\mathbf{b}\cdot\nabla_h u_h^N + cu_h^N, \phi - \mathcal{J}\phi)\end{aligned}$$

and

$$(\mathrm{curl}\psi, \nabla_h(u - u_h^N)) = (\mathrm{curl}(\psi - \mathcal{J}\psi), \nabla_h(u - u_h^N)) = \sum_{E\in\mathcal{E}_h} \int_E [\frac{\partial u_h^N}{\partial s}](\psi - \mathcal{J}\psi),$$

it follows that

$$\begin{aligned}&\mathcal{B}_h(u - u_h^N, u - u_h^N) = (R(u_h^N), \phi - \mathcal{J}\phi) + \sum_{E\in\mathcal{E}_h^0} \int_E [\mathbf{A}\nabla_h u_h^N \cdot \nu_E](\phi - \mathcal{J}\phi)ds \\ &+ \sum_{E\in\mathcal{E}_h} \int_E [\frac{\partial u_h^N}{\partial s}](\psi - \mathcal{J}\psi) + (\mathbf{b}\cdot\nabla_h(u - u_h^N) + c(u - u_h^N), u - u_h^N - \phi).\end{aligned}$$

Therefore, in view of Lemma 3.1

$$\begin{aligned}\mathcal{B}_h(u-u_h^N,u-u_h^N) \lesssim \mathrm{osc}_h|\phi|_{1,\Omega} + \eta_\nu(u_h^N)|\phi|_{1,\Omega} + \eta_s(u_h^N)|\psi|_{1,\Omega}\\ +(\mathbf{b}\cdot\nabla_h(u-u_h^N)+c(u-u_h^N),u-u_h^N-\phi).\end{aligned}$$

Taking advantage of

$$|\phi|_{1,\Omega} \lesssim ||\mathbf{A}^{\frac{1}{2}}\nabla\phi||_{0,\Omega} \quad , \quad |\psi|_{1,\Omega} \lesssim ||\mathbf{A}^{-\frac{1}{2}}\mathrm{curl}\psi||_{0,\Omega}$$

and

$$||\mathbf{A}^{\frac{1}{2}}\nabla\phi||_{0,\Omega}^2 + ||\mathbf{A}^{-\frac{1}{2}}\mathrm{curl}\psi||_{0,\Omega}^2 = (\mathbf{A}\nabla_h(u-u_h^N),\nabla_h(u-u_h^N)),$$

we get

$$\begin{aligned}&\mathcal{B}_h(u-u_h^N,u-u_h^N) \lesssim (\mathbf{A}\nabla_h(u-u_h^N),\nabla_h(u-u_h^N))^{\frac{1}{2}}(\mathrm{osc}_h + \eta_\nu(u_h^N)+\eta_s(u_h^N))\\ &+(||\mathbf{b}\cdot\nabla_h(u-u_h^N)||_{0,h} + c||u-u_h^N||_{0,\Omega})||u-u_h^N-\phi||_{0,\Omega}. \qquad (3.12)\end{aligned}$$

By assumptions **(b)** and **(c)**

$$||\mathbf{b}\cdot\nabla_h(u-u_h^N)||_{0,h} + ||c(u-u_h^N)||_{0,\Omega} \lesssim |||u-u_h^N|||_h,$$

which, together with (3.12), allows to deduce

$$\begin{aligned}\mathcal{B}_h(u-u_h^N,u-u_h^N) \lesssim |||u-u_h^N|||_h(\mathrm{osc}_h + \eta_\nu(u_h^N)+\eta_s(u_h^N)\\ +||u-u_h^N-\phi||_{0,\Omega}). \qquad (3.13)\end{aligned}$$

For the estimation of $||u-u_h^N-\phi||_{0,\Omega}$ we consider the following auxiliary problem

$$\begin{cases} -\Delta w_g = g, & \text{in } \Omega,\\ \quad w_g = 0, & \text{on } \partial\Omega. \end{cases} \qquad (3.14)$$

Green's formula implies

$$\begin{aligned}&(g,u-u_h^N-\phi) = (-\Delta w_g, u-u_h^N-\phi)\\ &= \sum_{T\in\mathcal{T}_h}\int_T \nabla w_g\cdot\nabla_h(u-u_h^N-\phi) - \int_{\partial T}\frac{\partial w_g}{\partial\nu}(u-u_h^N-\phi).\end{aligned}$$

Taking the regularity $w_g \in H^{1+\epsilon}(\Omega), \varepsilon \in (0,1]$, into account, by the standard estimation of the consistency error we obtain

$$\sum_{T\in\mathcal{T}_h}\int_{\partial T}\frac{\partial w_g}{\partial\nu}(u-u_h^N-\phi) \lesssim |w_g|_{1+\epsilon}(\sum_{T\in\mathcal{T}_h}h_T^{2\epsilon}|u-u_h^N-\phi|_{1,T}^2)^{\frac{1}{2}}.$$

Therefore, it follows that

$$\begin{aligned}(g,u-u_h^N-\phi) &\lesssim |w_g|_{1,\Omega}|u-u_h^N-\phi|_{1,h} + |w_g|_{1+\epsilon}(\sum_{T\in\mathcal{T}_h}h_T^{2\epsilon}|u-u_h^N-\phi|_{1,T}^2)^{\frac{1}{2}}\\ &\lesssim ||g||_{-1}|u-u_h^N-\phi|_{1,h} + ||g||_{\epsilon-1}|u-u_h^N-\phi|_{1,h}.\end{aligned}$$

Since $L^2(\Omega) \hookrightarrow H^{-1}(\Omega)$ and $L^2(\Omega) \hookrightarrow H^{\epsilon-1}(\Omega)$, we have $||g||_{-1} \lesssim ||g||_0$ and $||g||_{\epsilon-1} \lesssim ||g||_0$, whence

$$||u - u_h^N - \phi||_{0,\Omega} \lesssim |u - u_h^N - \phi|_{1,h}.$$

In view of Lemma 3.2 we obtain

$$||u - u_h^N - \phi||_{0,\Omega} \lesssim |u - u_h^N - \phi|_{1,h} \lesssim ||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega}. \tag{3.15}$$

Observing

$$\sum_{E\in\mathcal{E}_h} \int_E [\frac{\partial \phi}{\partial s}]\psi ds = 0,$$

we have

$$\begin{aligned}
||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega}^2 &= \sum_{T\in\mathcal{T}_h} (\mathbf{A}^{-1}\text{curl}\psi, \text{curl}\psi)_{0,T} \\
&= \sum_{T\in\mathcal{T}_h} (\nabla_h(u - u_h^N) - \nabla\phi, \text{curl}\psi)_{0,T} \\
&= \sum_{T\in\mathcal{T}_h} (\nabla_h(u - u_h^N), \text{curl}\psi)_{0,T} + \sum_{E\in\mathcal{E}_h} \int_E [\frac{\partial \phi}{\partial s}]\psi ds \\
&= \sum_{T\in\mathcal{T}_h} (\nabla_h(u - u_h^N), \text{curl}\psi)_{0,T}. \qquad (3.16)
\end{aligned}$$

For all $w \in V$ there holds

$$\sum_{T\in\mathcal{T}_h} (\nabla_h(u - w), \text{curl}\psi)_{0,T} = \sum_{E\in\mathcal{E}_h} \int_E [\frac{\partial(u - w)}{\partial s}]\psi ds = 0,$$

which, together with (3.16), yields

$$||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega}^2 = \sum_{T\in\mathcal{T}_h} (\nabla_h(w - u_h^N), \text{curl}\psi)_{0,T}.$$

Since

$$\begin{aligned}
||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega}^2 &= \sum_{T\in\mathcal{T}_h} (\mathbf{A}^{\frac{1}{2}}\nabla_h(w - u_h^N), \mathbf{A}^{-\frac{1}{2}}\text{curl}\psi)_{0,T} \\
&\leq ||\mathbf{A}^{\frac{1}{2}}\nabla_h(w - u_h^N)||_{0,h}||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega},
\end{aligned}$$

it follows that

$$||\mathbf{A}^{-\frac{1}{2}}\text{curl}\psi||_{0,\Omega} \leq ||\mathbf{A}^{\frac{1}{2}}\nabla_h(w - u_h^N)||_{0,h}. \tag{3.17}$$

Combining (3.15), (3.17) with Lemma 3.3, for the term $||u - u_h^N - \phi||_{0,\Omega}$ we finally obtain the following upper bound

$$||u - u_h^N - \phi||_{0,\Omega} \lesssim \eta_s(u_h^N). \tag{3.18}$$

Due to (3.13), (3.18) and Gårding's inequality (2.8) there holds

$$|||u - u_h^N|||_h^2 \lesssim |||u - u_h^N|||_h(\text{osc}_h + \eta_\nu(u_h^N) + \eta_s(u_h^N)) + \lambda||u - u_h^N||_{0,\Omega}^2. \tag{3.19}$$

Since $||\cdot||_{1,h}$ and $|||\cdot|||_h$ are equivalent norms, we have

$$C_1||v||_{1,h}^2 \le |||v|||_h^2 \le C_2||v||_{1,h}^2 \quad , \quad v \in W = V \oplus V_h^N. \tag{3.20}$$

Moreover, by the Aubin-Nitsche duality technique [20] we can easily show

$$||u - u_h^N||_{0,\Omega} \le C_3 h^\alpha ||u - u_h^N||_{1,h}, \tag{3.21}$$

where $\alpha \in (0,1]$ is from (2.4). Taking (3.19), (3.20) and (3.21) into account , for $h_0^{2\alpha} < C_1(\lambda C_3^2)^{-1}$ the assertion (3.11) holds true. □

3.2. Efficiency of a posteriori error estimators. This section is devoted to the efficiency of the a posteriori error estimator which will be established by a series of lemmas.

LEMMA 3.4. *For the edge residual $\eta_{s,E}(u_h^N), E \in \mathcal{E}_h$, given by* (3.1) *there holds*

$$\eta_{s,E}(u_h^N) \lesssim ||\nabla_h(u - u_h^N)||_{0,w_E} \quad , \quad E \in \mathcal{E}_h. \tag{3.22}$$

Proof. Denoting by b_E the edge bubble function, we have

$$\eta_{s,E}^2(u_h^N) = \frac{3h_E}{2} \int_E [\frac{\partial u_h^N}{\partial s}][\frac{\partial u_h^N}{\partial s}] b_E.$$

Setting $\psi_E = [\frac{\partial u_h^N}{\partial s}] b_E$, it follows that

$$\eta_{s,E}^2(u_h^N) = \frac{3h_E}{2}(\nabla_h(u - u_h^N), \text{curl}\psi_E)_{0,w_E} \lesssim h_E||\nabla_h(u - u_h^N)||_{0,w_E}||\text{curl}\psi_E||_{0,w_E}.$$

Observing

$$||\text{curl}\psi_E||_{0,w_E} \lesssim h_E^{-1}||\psi_E||_{0,w_E} \lesssim h^{-\frac{1}{2}}||[\frac{\partial u_h^N}{\partial s}]||_{0,E}$$

and the previous inequality allows to conclude. □

LEMMA 3.5 (Discrete local efficiency [9]). *For any refined edge $E \in \mathcal{E}_H$ there holds*

$$\eta_{s,E}(u_H^N) \lesssim ||\nabla_h(u_h^N - u_H^N)||_{0,w_E} \quad , \quad E \in \mathcal{E}_H. \tag{3.23}$$

LEMMA 3.6. *For the edge residual $\eta_{\nu,E}(u_h^n)$ given by* (3.3) *there holds*

$$\eta_{\nu,E}(u_h^N) \lesssim \text{osc}_h(w_E) + J_h(w_E). \tag{3.24}$$

Proof. Observing that $[\nabla u_h^N] \cdot \nu_E$ is constant on E and denoting by $\varphi_E \in V_h^N$ the edge basis function with supp$\varphi_E \subset \overline{w_E}$, we have

$$||[\mathbf{A}\nabla u_h^N] \cdot \nu_E||_{0,E}^2 \lesssim ([\nabla u_h^N] \cdot \nu_E, \varphi_E[\nabla u_h^N] \cdot \nu_E)_{0,E}$$

By the assumption **(a)** on $\mathbf{A}$ we obtain

$$([\nabla u_h^N] \cdot \nu_E, \varphi_E)_{0,E} \lesssim |([\mathbf{A}\nabla u_h^N] \cdot \nu_E, \varphi_E)_{0,E}|,$$

and

$$
\begin{aligned}
([\mathbf{A}\nabla u_h^N]\cdot\nu_E,\varphi_E)_{0,E} &= ([\mathbf{A}\nabla u_h^N]\cdot\nu_E,\varphi_E)_{0,E} + ([\mathbf{A}\nabla u_h^N]\cdot\nu_E,\varphi_E)_{0,w_E}\\
&= (\mathbf{A}\nabla u_h^N,\nabla\varphi_E)_{0,w_E} + (\mathrm{div}(\mathbf{A}\nabla u_h^N),\varphi_E)_{0,w_E}\\
&= (R_{w_E}(u_h^N),\varphi_E)_{0,w_E}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\eta_{\nu,E}^2(u_h^N) &\lesssim h_E([\nabla u_h^N]\cdot\nu_E,\varphi_E[\nabla u_h^N]\cdot\nu_E)_{0,E}\\
&\lesssim \eta_{\nu,E}(u_h^N)|(R_{w_E}(u_h^N),\varphi_E)_{0,w_E}|\\
&\lesssim \eta_{\nu,E}(u_h^N)|(R_{w_E}(u_h^N)-\overline{R_{w_E}(u_h^N)},\varphi_E)_{0,w_E}\\
&\quad+(\overline{R_{w_E}(u_h^N)},\varphi_E)_{0,w_E}|. \qquad (3.25)
\end{aligned}
$$

Then, (3.24) is a direct consequence of (3.25). □

Remark: In view of Lemma 3.6, we easily deduce the following reliability result which is an improvement of Theorem 3.1,

$$
|||u-u_h^N|||_h \lesssim \eta_s(u_h^N)+J_h.
$$

LEMMA 3.7. *For the element residuals J_T and J_{ω_E} as given by (3.5) there holds*

$$
J_T \lesssim ||u-u_h^N||_{1,T}+\mathrm{osc}_h(T), \qquad (3.26)
$$

$$
J_{w_E} \lesssim \mathrm{osc}_h(w_E)+\eta_{\nu,E}(u_h^N). \qquad (3.27)
$$

Proof. We set $\phi_T = b_T\overline{R_T(u_h^N)}$, where b_T is the bubble function on T with $b_T(x_T)=1$. Taking advantage of

$$
\begin{aligned}
h_T^2\int_T R_T(u_h^N)\phi_T dx &= h_T^2\mathcal{B}_h(u-u_h^N,\phi_T)\\
&\lesssim h_T^2||u-u_h^N||_{1,T}||\phi_T||_{1,T}\\
&\lesssim \frac{1}{\delta}||u-u_h^N||_{1,T}^2+\delta h_T^2||\overline{R_T(u_h^N)}||_{0,T}^2,
\end{aligned}
$$

we obtain

$$
\begin{aligned}
h_T^2\int_T \overline{R_T(u_h^N)}^2 dx &= \frac{20}{9}h_T^2\int_T \overline{R_T(u_h^N)}\phi_T dx\\
&= \frac{20}{9}h_T^2\int_T (\overline{R_T(u_h^N)}-R_T(u_h^N)+R_T(u_h^N))\phi_T dx\\
&\lesssim \mathrm{osc}_h^2(T)+\frac{1}{\delta}||u-u_h^N||_{1,T}^2+\delta h_T^2||\overline{R_T(u_h^N)}||_{0,T}^2. \qquad (3.28)
\end{aligned}
$$

Choosing δ in (3.28) appropriately, it follows that

$$
h_T^2\int_T \overline{R_T(u_h^N)}^2 dx \lesssim ||u-u_h^N||_{1,T}^2+\mathrm{osc}_h^2(T), \qquad (3.29)
$$

from which we easily deduce (3.26).

For the proof of (3.27) we note that

$$
h_E||\overline{R_{w_E}(u_h^N)}||_{0,w_E} \lesssim |w_E||\overline{R_{w_E}(u_h^N)}|,
$$

and

$$
\begin{aligned}
&|w_E||\overline{R_{w_E}(u_h^N)}| = 3|(\overline{R_{w_E}(u_h^N)}, \varphi_E)| \\
&\leq 3|(\overline{R_{w_E}(u_h^N)} - R_{w_E}(u_h^N), \varphi_E)_{0,w_E}| + 3|(R_{w_E}(u_h^N), \varphi_E)_{0,w_E}|,
\end{aligned}
$$

where φ_E is the edge basis function in V_h^N. Obviously, we have

$$
(R_{w_E}(u_h^N), \varphi_E))_{0,w_E} = \int_E [\mathbf{A}\nabla u_h^N] \cdot \nu_E \varphi_E ds \lesssim \eta_{\nu,E}.
$$

Combining the previous inequalities results in (3.27). ∎

4. Quasi-Orthogonality. In this section, we generalize the result in [9] to non-symmetric and indefinite problems. The main difficulty is how to deal with the non-symmetric and indefinite bilinear form. Here, we require $\mathbf{A}$ to be piecewise constant.

LEMMA 4.1. *Let $\varepsilon_H = u_h^N - u_H^N$ and $e_h = u - u_h^N$. Then there holds*

$$
\mathcal{B}_h(\varepsilon_H, e_h) \leq C_0(J_h + J_H)(|||u - u_h^N|||_h + |||u - u_H^N|||_H). \tag{4.1}
$$

Proof. The original problem (2.5) can be rewritten according to

$$
(\mathbf{A}\nabla_h u_h^N, \nabla_h v_h^N) = (R(u_h^N), v_h^N) \quad , \quad v_h^N \in V_h^N. \tag{4.2}
$$

where $R(u_h^N)$ is given by

$$
R(u_h^N) = f - \mathbf{b} \cdot \nabla_h u_h^N - c u_h^N,
$$

We consider the auxiliary problem

$$
(\mathbf{A}\nabla z, \nabla v) = (\overline{R(u_h^N)}, v) \quad , \quad v \in V. \tag{4.3}
$$

The nonconforming finite element approximation of (4.3) requires the computation of $z_h^N \in V_h^N$ such that

$$
(\mathbf{A}\nabla_h z_h^N, \nabla_h v_h^N) = (\overline{R(u_h^N)}, v_h^N), \qquad \forall v_h^N \in V_h^N. \tag{4.4}
$$

The mixed formulation of (4.3) involves the flux $\sigma = \mathbf{A}\nabla z$. We refer to $Q_h := RT_0(\Omega, \mathcal{T}_h)$ as the lowest order Raviart-Thomas finite element space and to W_h as the linear space of elementwise constants with respect to the triangulation $\mathcal{T}_h$. Then, the lowest order Raviart-Thomas mixed finite element approximation of (4.3) amounts to the computation of $\sigma_h^M \in Q_h$ and $u_h^M \in W_h$ such that

$$
\left\{
\begin{aligned}
&(\mathbf{A}^{-1}\sigma_h^M, q_h^M) + (u_h^M, \mathrm{div} q_h^M) = 0 \ , \ q_h^M \in Q_h, \\
&(\mathrm{div}\sigma_h^M, v_h^M) = -(\overline{R(u_h^N)}, v_h^M) \qquad , \ v_h^M \in W_h.
\end{aligned}
\right. \tag{4.5}
$$

Applying the techniques used to prove the equivalence of the nonconforming and the lowest order Raviart-Thomas mixed finite element approximation for Poisson's equation (cf., e.g., [4]), we get the following relationship between the solutions of (4.4) and (4.5)

$$
\sigma_h^M|_T = \mathbf{A}\nabla_h z_h^N|_T - \frac{1}{2}\overline{R_T(u_h^N)}(x - x_T) \quad , \quad T \in \mathcal{T}_h. \tag{4.6}
$$

Therefore

$$\begin{aligned}\mathcal{B}_h(\varepsilon_H, e_h) &= (\mathbf{A}\nabla_h u_h^N - \mathbf{A}\nabla_h z_h^N + \mathbf{A}\nabla_h z_h^N - \sigma_h^M + \sigma_h^M - \sigma_H^M + \sigma_H^M - \mathbf{A}\nabla_H z_H^N \\ &+\mathbf{A}\nabla_H z_H^N - \mathbf{A}\nabla_H u_H^N, \nabla_h(u-u_h^N)) + (R(u_H^N), u-u_h^N) - (R(u_h^N), u-u_h^N). \quad (4.7)\end{aligned}$$

Subtracting (4.4) from (4.2) gives

$$\begin{aligned}(\mathbf{A}\nabla_h(u_h^N - z_h^N), \nabla_h v_h^N) &= (R(u_h^N) - \overline{R(u_h^N)}, v_h^N) \\ &= (R(u_h^N) - \overline{R(u_h^N)}, v_h^N - \mathcal{P}_0 v_h^N), \quad (4.8)\end{aligned}$$

where $\mathcal{P}_0 : V_h^N \to W_h$ stands for the L^2 projection. Setting $v_h^N = u_h^N - z_h^N$ in (4.8) yields

$$||\mathbf{A}^{\frac{1}{2}}\nabla_h(u_h^N - z_h^N)||_{0,h} \lesssim \mathrm{osc}_h. \quad (4.9)$$

From equality (4.6) we deduce

$$(\mathbf{A}\nabla_h z_h^N - \sigma_h^M, \nabla_h(u-u_h^N)) \lesssim J_h|u-u_h^N|_{1,h}. \quad (4.10)$$

Similarly, on the coarse level $\mathcal{T}_H$, we also have

$$||\mathbf{A}^{\frac{1}{2}}\nabla_H(z_H^N - u_H^N)||_{0,H} \lesssim \mathrm{osc}_H, \quad (4.11)$$

$$(\sigma_H^M - \mathbf{A}\nabla_H z_H^N, \nabla_h(u-u_h^N)) \lesssim J_H|u-u_h^N|_{1,h}. \quad (4.12)$$

Making use of (4.5), we get

$$\mathrm{div}\sigma_h^M = -\overline{R(u_h^N)} \quad , \quad \mathrm{div}\sigma_H^M = -\overline{R(u_H^N)}.$$

It is easy to see that $(\sigma_h^M - \sigma_H^M)\cdot\nu_E$ is constant on $E \in \mathcal{E}_h^0$ and $\int_E [u-u_h^N]ds = 0$. Consequently, there holds

$$(\sigma_h^M - \sigma_H^M, \nabla_h(u-u_h^N)) = (u-u_h^N, \overline{R(u_h^N)} - \overline{R(u_H^N)}). \quad (4.13)$$

Combining (4.7)-(4.13) implies

$$\begin{aligned}\mathcal{B}_h(\varepsilon_H, e_h) &\lesssim (\mathrm{osc}_h + \mathrm{osc}_H)||\mathbf{A}^{\frac{1}{2}}\nabla_h(u-u_h^N)||_{0,h} + (J_h + J_H)|u-u_h^N|_{1,h} \\ &+(u-u_h^N, R(u_H^N) - \overline{R(u_H^N)}) - (u-u_h^N, R(u_h^N) - \overline{R(u_h^N)}). \quad (4.14)\end{aligned}$$

Let $(e_H)_T$ and $(\varepsilon_H)_T$ be the mean values of $e_H|_T$ and $\varepsilon_H|_T$ with respect to T. Then, there holds

$$\begin{aligned}(u-u_h^N, R(u_H^N) - \overline{R(u_H^N)})_{0,T} &= (e_H, R(u_H^N) - \overline{R(u_H^N)})_{0,T} - (\varepsilon_H, R(u_H^N) - \overline{R(u_H^N)})_{0,T} \\ &= (e_H - (e_H)_T, R(u_H^N) - \overline{R(u_H^N)})_{0,T} \\ &\quad -(\varepsilon_H - (\varepsilon_H)_T, R(u_H^N) - \overline{R(u_H^N)})_{0,T} \\ &\lesssim \mathrm{osc}_T(u_H^N)(||\nabla_H(u-u_H^N)||_{0,T} + ||\nabla_h(u_h^N - u_H^N)||_{0,T}) \\ &\lesssim \mathrm{osc}_T(u_H^N)(||\nabla_H(u-u_H^N)||_{0,T} + ||\nabla_h(u-u_h^N)||_{0,T}).\end{aligned}$$

Similarly, for $T' \in \mathcal{T}_h$ we have

$$\begin{aligned}(u-u_h^N, R(u_h^N) - \overline{R(u_h^N)})_{0,T'} &= (e_h - (e_h)_{T'}, R(u_h^N) - \overline{R(u_h^N)})_{0,T'} \\ &\lesssim \mathrm{osc}_{T'}(u_h^N)||\nabla_h(u-u_h^N)||_{0,T'}.\end{aligned}$$

Finally, we arrive at

$$\begin{aligned}\mathcal{B}_h(\varepsilon_H, e_h) &\lesssim (\mathrm{osc}_h + \mathrm{osc}_H + J_h + J_H)(|||u - u_h^N|||_h + |||u - u_H^N|||_H) \\ &\lesssim (J_h + J_H)(|||u - u_h^N|||_h + |||u - u_H^N|||_H),\end{aligned}$$

which completes the proof. $\square$

The following theorem highlights the relationship between e_h and ε_H.

THEOREM 4.1 (Quasi-orthogonality). *Let h_0 be the initial mesh size and let α be the parameter in the regularity assumption* (2.4). *Then, there holds*

$$|||e_h|||_h^2 \leq \Lambda_0 |||e_H|||_H^2 - |||\varepsilon_H|||_h^2 + \Lambda_1 (J_h^2 + J_H^2), \tag{4.15}$$

where

$$\Lambda_0 = \frac{1 + O(h_0^\alpha)}{1 - O(h_0^\alpha)}$$

and Λ_1 depends on the data of the problem.

Proof. By Gårding's inequality (2.8) we have

$$|||e_h|||_h^2 \leq \mathcal{B}_h(e_h, e_h) + \lambda ||e_h||_{0,\Omega}^2, \tag{4.16}$$

$$B_H(\varepsilon_H, \varepsilon_H) \geq |||\varepsilon_H|||_h^2 - \lambda ||\varepsilon_H||_{0,\Omega}^2. \tag{4.17}$$

A simple calculation shows

$$\begin{aligned}\mathcal{B}_h(e_h, e_h) = \mathcal{B}_H(e_H, e_H) - \mathcal{B}_h(\varepsilon_H, \varepsilon_H) - 2\mathcal{B}_h(\varepsilon_H, e_h) \\ +(\mathbf{b} \cdot \nabla_h \varepsilon_H, e_h) - (\mathbf{b} \cdot \nabla_h e_h, \varepsilon_H).\end{aligned} \tag{4.18}$$

In view of (3.20) and (3.21) it is easy to check that

$$\begin{aligned}(\mathbf{b} \cdot \nabla_h \varepsilon_H, e_h) &\leq B ||\varepsilon_H||_{1,h} C_3 h^\alpha ||e_h||_{1,h} \\ &\leq \frac{1}{2C_1}(B^2 C_3^2 h^{2\alpha} \delta |||e_h|||_h^2 + \frac{1}{\delta} |||\varepsilon_H|||_h^2)\end{aligned} \tag{4.19}$$

and

$$(\mathbf{b} \cdot \nabla_h e_h, \varepsilon_H) \leq \frac{B C_3 h^\alpha}{C_1} |||e_h|||_h^2 + \frac{1}{2C_1}(\frac{B^2}{\delta} |||e_H|||_H^2 + C_3^2 H^{2\alpha} \delta |||e_h|||_h^2), \tag{4.20}$$

where $B = ||\mathbf{b}||_{0,\infty}$ and $\delta > 0$ is a constant that will be chosen later. Moreover, with the same δ as above we have

$$\mathcal{B}_H(e_H, e_H) \leq (1 + \frac{B}{2C_1}(C_3^2 h^{2\alpha} \delta + \frac{1}{\delta})) |||e_H|||_H^2. \tag{4.21}$$

Combining (4.16)-(4.21) and using Lemma 4.1 yield

$$\begin{aligned}|||e_h|||_h^2 \leq\ & (1 + \frac{B}{2C_1}(C_3^2 h^{2\alpha} \delta + \frac{1}{\delta})) |||e_H|||_H^2 - |||\varepsilon_H|||_h^2 + \lambda ||\varepsilon_H||_{0,\Omega}^2 \\ & + 2C_0 (J_h + J_H)(|||e_h|||_h + |||e_H|||_H) \\ & + \frac{1}{2C_1}(B^2 C_3^2 h^{2\alpha} \delta |||e_h|||_h^2 + \frac{1}{\delta} |||\varepsilon_H|||_h^2) \\ & + \frac{B C_3 h^\alpha}{C_1} |||e_h|||_h^2 + \frac{1}{2C_1}(\frac{B^2}{\delta} |||e_H|||_H^2 + C_3^2 H^{2\alpha} \delta |||e_h|||_h^2) \\ & + \frac{\lambda C_3^2 h^{2\alpha}}{C_1} |||e_h|||_h^2.\end{aligned} \tag{4.22}$$

An application of the Aubin-Nitsche duality technique and Young's inequality gives

$$\lambda||\varepsilon_H||^2_{0,\Omega} \leq \lambda(||e_h||^2_{0,\Omega} + ||e_H||^2_{0,\Omega})$$
$$\leq \frac{\lambda C_3^2}{C_1}(h^{2\alpha}|||e_h|||^2_h + H^{2\alpha}|||e_H|||^2_H),$$

and

$$2C_0(J_h + J_H)(|||e_h|||_h + |||e_H|||_H) \leq \frac{4C_0^2}{\xi}(J_h^2 + J_H^2) + \xi(|||e_h|||^2_h + |||e_H|||^2_H),$$

where $h, H < h_0$ and $\xi > 0$ will be specified below.

It follows from the previous two inequalities and (4.22) that

$$\begin{aligned}
&(1 - \frac{2\lambda C_3^2 h_0^{2\alpha}}{C_1} - \xi - \frac{C_3^2 B^2 h_0^{2\alpha}\delta}{2C_1} - \frac{BC_3 h_0^\alpha}{C_1} - \frac{C_3^2 h_0^{2\alpha}\delta}{2C_1})|||e_h|||^2_h \\
&\leq (1 + \frac{B}{2C_1}(C_3^2 h_0^{2\alpha}\delta + \frac{1}{\delta}) + \frac{\lambda C_3^2 h_0^{2\alpha}}{C_1} + \xi + \frac{B^2}{2C_1\delta})|||e_H|||^2_H \\
&\quad -(1 - \frac{1}{2C_1\delta})|||\varepsilon_H|||^2_h + \frac{4C_0^2}{\xi}(J_h^2 + J_H^2). \qquad (4.23)
\end{aligned}$$

Choosing $\xi = h_0^\alpha$, in view of

$$1 - \frac{2\lambda C_3^2 h_0^{2\alpha}}{C_1} - \xi - \frac{C_3^2 B^2 h_0^{2\alpha}\delta}{2C_1} - \frac{BC_3 h_0^\alpha}{C_1} - \frac{C_3^2 h_0^{2\alpha}\delta}{2C_1} = 1 - \frac{1}{2C_1\delta}$$

we easily get

$$\delta = \frac{-(4\lambda C_3^2 h_0^\alpha + 2C_1 + 2BC_3) + \sqrt{(4\lambda C_3^2 h_0^\alpha + 2C_1 + 2BC_3)^2 + 4C_3^2(B^2+1)}}{2C_3^2(B^2+1)h_0^\alpha},$$

i.e., $\delta \approx h_0^{-\alpha}$. Hence, if h_0 is sufficiently small, then $(2C_1\delta)^{-1} < 1$.

Setting

$$\Lambda_h = 1 - \frac{2\lambda C_3^2 h_0^{2\alpha}}{C_1} - \xi - \frac{C_3^2 B^2 h_0^{2\alpha}\delta}{2C_1} - \frac{BC_3 h_0^\alpha}{C_1} - \frac{C_3^2 h_0^{2\alpha}\delta}{2C_1}$$

and

$$\Lambda_H = 1 + \frac{B}{2C_1}(C_3^2 h_0^{2\alpha}\delta + \frac{1}{\delta}) + \frac{\lambda C_3^2 h_0^{2\alpha}}{C_1} + \xi + \frac{B^2}{2C_1\delta},$$

implies

$$\Lambda_h = 1 - O(h_0^\alpha) \quad , \quad \Lambda_H = 1 + O(h_0^\alpha).$$

We set $\Lambda_0 = \Lambda_H/\Lambda_h$ and $\Lambda_1 = 4C_0^2 h_0^{-\alpha}/\Lambda_h$. Then, (4.15) is a consequence of (4.23). □

5. ANFEM I and its convergence analysis.

5.1. The algorithm ANFEM I. The ANFEM procedure consists of adaptive loops of the cycle (1.1). In the step SOLVE, we choose some well-known iterative schemes such as GMRES to solve (2.5). In the step ESTIMATE, we adopt the reliable and efficient a posteriori error estimator suggested in section 3. As far as the step REFINE is concerned, we apply bisection (without the interior node property introduced in [9]). Since the step MARK plays a crucial role in the ANFEM algorithm, this section mainly addresses the implementation of this step:

MARK algorithm:

Let the set $\mathcal{M}$ consist of marked edges and marked triangles, i.e., $\mathcal{M} = \{\mathcal{F}, \hat{\mathcal{T}}\}$, where $\mathcal{F} \subset \mathcal{E}_H$ and $\hat{\mathcal{T}} \subset \mathcal{T}_H$.
1. Given parameters $0 < \theta_1, \theta_2 < 1$, set $\mathcal{M}_0 = \emptyset$;
2. Mark a set of edges $\mathcal{F} \subset \mathcal{E}_H$ with minimal cardinality such that

$$\theta_1 \eta_s^2(u_H^N) \leq \eta_s^2(u_H^N, \mathcal{F}). \tag{5.1}$$

Set $\mathcal{M} = \mathcal{M}_0 \cup \{\mathcal{F}\}$;
3. Mark a set of triangles $\hat{\mathcal{T}} \subset \mathcal{T}_H$ with minimal cardinality such that

$$\theta_2 J_H^2 \leq J_H^2(\hat{\mathcal{T}}). \tag{5.2}$$

Set $\mathcal{M} = \mathcal{M} \cup \{\hat{\mathcal{T}}\}$;
4. Mark further edges and triangles to avoid hanging nodes in order to maintain the shape regularity of the mesh in the refinement step;
5. Output the marked edges and triangles $\mathcal{M}$:=MARK$(\mathcal{E}_H, \mathcal{T}_H)$.

Based on the step MARK, the algorithm ANFEM I is given as follows:

Algorithm ANFEM I:
1. Given parameters $0 < \theta_1, \theta_2 < 1$ and an initial mesh $\mathcal{T}_0$, set $k = 0$;
2. Solve system (2.5) and get u_k^N;
3. Compute $\eta_{k,s,E}$ on each $E \subset \mathcal{E}_k$, and $J_{k,T}$ on each $T \subset \mathcal{T}_k$;
4. Implement the step MARK and get $\mathcal{M}_k$ containing the marked edges and triangles;
5. Refine the mesh $\mathcal{T}_k$;
6. Set $k = k + 1$ and go to step 2.

5.2. Convergence analysis. We are now in a position to prove the convergence result. As a prerequisite, we provide the following lemma.

LEMMA 5.1. *Suppose that* (5.2) *is implemented in the step* MARK. *Then, there exist constants* $0 < \rho_1 < 1$ *and* $C_R = O(h_0^2)$ *such that*

$$J_h^2 \leq \rho_1 J_H^2 + C_R |||\varepsilon_H|||_h^2. \tag{5.3}$$

Proof. For $T \in \mathcal{T}_h$ and $T' \in \mathcal{T}_H$ we set $h_T = \gamma_{T'} h_{T'}$ where

$$\gamma_{T'} = \begin{cases} \gamma_0 \ , \ T' \in \widehat{\mathcal{T}_H} \ , \ 0 < \gamma_0 < 1 \\ 1 \ , \ T' \notin \widehat{\mathcal{T}_H} \end{cases} .$$

Then

$$\sum_{T\in\mathcal{T}_h(T')} J_T^2(u_h^N) \le \sum_{T\in\mathcal{T}_h(T')} (1+\delta_0)h_T^2||R_T(u_H^N)||_{0,T}^2 + (1+\delta_0^{-1})h_T^2||\mathcal{L}_T(\varepsilon_H)||_{0,T}^2$$
$$= (1+\delta_0)\gamma_{T'}^2 J_{T'}^2 + (1+\delta_0^{-1}) \sum_{T\in\mathcal{T}_h(T')} h_T^2||\mathcal{L}_T(\varepsilon_H)||_{0,T}^2,$$

where $\mathcal{L}_T(\varepsilon_H) = (\mathbf{b}\cdot\nabla_h\varepsilon_H + c\varepsilon_H)|_T$. By (5.2) we have

$$\begin{aligned}\sum_{T'\in\mathcal{T}_H} \gamma_{T'}^2 J_{T'}^2 &= \gamma_0^2 \sum_{T'\in\widehat{\mathcal{T}_H}} J_{T'}^2 + \sum_{T'\in\mathcal{T}_H\setminus\widehat{\mathcal{T}_H}} J_{T'}^2\\ &= J_H^2 - (1-\gamma_0^2)\sum_{T'\in\widehat{\mathcal{T}_H}} J_{T'}^2\\ &\le (1-(1-\gamma_0^2)\theta_2)J_H^2,\end{aligned}$$

and hence, for some $\delta_0 > 0$

$$J_h^2 \le (1+\delta_0)(1-(1-\gamma_0^2)\theta_2)J_H^2 + (1+\delta_0^{-1})\sum_{T\in\mathcal{T}_h} h_T^2||\mathcal{L}_T(\varepsilon_H)||_{0,T}^2.$$

It is easy to see that

$$\sum_{T\in\mathcal{T}_h} h_T^2||\mathcal{L}_T(\varepsilon_H)||_{0,T}^2 \le C_4 h_0^2|||\varepsilon_H|||_h^2.$$

We choose δ_0 small enough such that

$$\rho_1 = (1+\delta_0)(1-(1-\gamma_0^2)\theta_2) < 1,$$

and choose C_R according to

$$C_R = C_4(1+\delta_0^{-1})h_0^2$$

which then proves (5.3). □

Theorem 5.1. *Under the assumption that the initial mesh size h_0 is sufficiently small there exist positive constants $0 < \rho < 1$ and $\beta > 0$ such that*

$$|||u-u_h^N|||_h^2 + \beta J_h^2 \le \rho(|||u-u_H^N|||_H^2 + \beta J_H^2). \tag{5.4}$$

Proof. In view of the quasi-orthogonality we have

$$\begin{aligned}|||e_h|||_h^2 &\le \Lambda_0|||e_H|||_H^2 - |||\varepsilon_H|||_h^2 + \Lambda_1(J_h^2+J_H^2)\\ &= \Lambda_0|||e_H|||_H^2 - \beta_0|||\varepsilon_H|||_h^2 - (1-\beta_0)|||\varepsilon_H|||_h^2 + \Lambda_1(J_h^2+J_H^2),\end{aligned} \tag{5.5}$$

where $0 < \beta_0 < 1$ will be chosen later. Based on the implementation of (5.1), it follows from Theorem 3.1 and Lemmas 3.5, 3.6 that

$$|||e_H|||_H^2 \lesssim |||\varepsilon_H|||_h^2 + J_H^2.$$

Consequently, there always exists $0 < \delta_0 < 1$ such that

$$\delta_0|||e_H|||_H^2 \le |||\varepsilon_H|||_h^2 + J_H^2,$$

whence

$$-\beta_0|||\varepsilon_H|||_h^2 \leq -\beta_0\delta_0|||e_H|||_H^2 + \beta_0 J_H^2. \tag{5.6}$$

Inserting (5.6) into (5.5) gives

$$|||e_h|||_h^2 \leq (\Lambda_0 - \beta_0\delta_0)|||e_H|||_H^2 + \beta_0 J_H^2 - (1-\beta_0)|||\varepsilon_H|||_h^2 + \Lambda_1(J_h^2 + J_H^2). \tag{5.7}$$

We claim that for sufficiently small h_0

$$0 < \Lambda_0 - \beta_0\delta_0 < 1. \tag{5.8}$$

A simple calculation shows that (5.8) holds true, if

$$O(h_0^\alpha) < \frac{\beta_0\delta_0}{2 + \beta_0\delta_0}$$

is satisfied. Observing (5.2), (5.7) and Lemma 5.1 results in

$$\begin{aligned} |||e_h|||_h^2 + \beta_1 J_h^2 \leq (\Lambda_0 - \beta_0\delta_0)|||e_H|||_H^2 + \beta_0 J_H^2 - (1-\beta_0)|||\varepsilon_H|||_h^2 + \\ +\Lambda_1(J_h^2 + J_H^2) + \beta_1\rho_1 J_H^2 + \beta_1 C_R|||\varepsilon_H|||_h^2, \end{aligned} \tag{5.9}$$

where $\beta_1 > 0$ will be determined later. We select β_0, β_1 in such a way that

$$1 - \beta_0 = \beta_1 C_R, \qquad \beta_0 + \beta_1\rho_1 + \wedge_1 = \beta_1\gamma_1 - \wedge_1\gamma_1,$$

where γ_1 satisfies $0 < \rho_1 < \gamma_1 < 1$. A simple calculation shows that

$$\beta_1 = \frac{1 + \Lambda_1 + \Lambda_1\gamma_1}{C_R + \gamma_1 - \rho_1}, \qquad \beta_0 = 1 - \frac{1 + \Lambda_1 + \Lambda_1\gamma_1}{C_R + \gamma_1 - \rho_1} C_R.$$

Consequently, $0 < \beta_0 < 1$ is guaranteed provided that

$$C_R < \frac{\gamma_1 - \rho_1}{\Lambda_1(1 + \gamma_1)},$$

which holds true for sufficiently small h_0. It is obvious that $\beta_1 > \Lambda_1$. In order to complete the proof, we set

$$0 < \rho = \max\{\Lambda_0 - \beta_0\delta_0, \gamma_1\} < 1 \quad , \quad \beta = \beta_1 - \Lambda_1 > 0,$$

which in combination with (5.9) proves (5.4). □

6. ANFEM II and its optimal complexity. In the algorithm ANFEM I, one indicator may be small compared to the other one which renders the step MARK as being not optimal. In this section, we suggest ANFEM II by a modification of the step MARK in ANFEM I and obtain optimal complexity. In the course of the analysis, we have a quasi-discrete reliability result.

As a prerequisite, we introduce two intergrid transfer operators I_h^H and I_H^h between two successive nonconforming spaces. The transfer operator I_h^H is defined by

$$\int_E I_h^H v_h^N ds = \int_E v_h^N ds, \quad \text{for} \quad \forall E \in \mathcal{E}_H \quad , \quad v_h^N \in V_h^N,$$

whereas I_H^h is given according to

$$\int_E I_H^h v_H ds = \begin{cases} \int_E v_H^N ds, & \text{for } E \in \mathcal{E}_h \bigcap \mathcal{E}_H \text{ or } E \\ & \text{in the interior of } T \in \mathcal{T}_H, \\ \\ \frac{1}{2}\int_E (v_H^N|_{T_1} + v_H^N|_{T_2}) ds, & \text{for } E \in \mathcal{E}_h \backslash \mathcal{E}_H \text{ and } E \text{ not} \\ & \text{in the interior of } T \in \mathcal{T}_H. \end{cases}$$

As has been shown in [30], there holds

$$||(I - I_h^H) v_h^N||_{0,T} \lesssim h_T ||\nabla_h v_h^N||_{0,h} \quad , \quad T \in \mathcal{T}_H. \tag{6.1}$$

Choosing $v_h^N = I_H^h u_H^N$ and taking advantage of (3.10), we deduce

$$\inf_{v_h^N \in V_h^N} ||\nabla_h (u_H^N - v_h^N)||_{0,h}^2 \lesssim \eta_s^2(u_H^N, \mathcal{F}), \tag{6.2}$$

where $\mathcal{F} \subset \mathcal{E}_H$ refers to the set of edges refined in the adaptive loop (1.1).

Lemma 6.1 (Quasi-discrete reliability). *There exists a constant $C_5 > 0$ such that*

$$\mathcal{B}_h(u_h^N - u_H^N, u_h^N - u_H^N) \le C_5 (J_H^2(\mathcal{M}) + \eta_s^2(u_H^N, \mathcal{F})), \tag{6.3}$$

where $\mathcal{M} \subset \mathcal{T}_H$ and $\mathcal{F} \subset \mathcal{E}_H$ are any subsets marked for refinement.

Proof. For any $v_h^N \in V_h^N$, we decompose the bilinear form by means of

$$\mathcal{B}_h(u_h^N - u_H^N, u_h^N - u_H^N) = \mathcal{B}_h(u_h^N - u_H^N, u_h^N - v_h^N) + \mathcal{B}_h(u_h^N - u_H^N, v_h^N - u_H^N). \tag{6.4}$$

For the estimation of the first term on the right hand side of (6.4), we note that

$$(\mathbf{A}\nabla_h(u_h^N - u_H^N), \nabla_h(u_h^N - v_h^N)) = (R(u_h^N), u_h^N - v_h^N) \\ -(R(u_H^N), I_h^H(u_h^N - v_h^N)) + (\mathbf{A}\nabla_H u_H^N, \nabla_H I_h^H(u_h^N - v_h^N) - \nabla_h(u_h^N - v_h^N)).$$

In view of the definition of I_h^H, an elementwise application of Green's formula yields

$$(\mathbf{A}\nabla_H u_H^N, \nabla_H I_h^H(u_h^N - v_h^N) - \nabla_h(u_h^N - v_h^N)) = 0,$$

whence

$$(\mathbf{A}\nabla_h(u_h^N - u_H^N), \nabla_h(u_h^N - v_h^N)) \\ = (R(u_h^N), (I - I_h^H)(u_h^N - v_h^N)) + (R(u_h^N) - R(u_H^N), I_h^H(u_h^N - v_h^N)).$$

We further get

$$\mathcal{B}_h(u_h^N - u_H^N, u_h^N - v_h^N) \\ = (R(u_h^N), (I - I_h^H)(u_h^N - v_h^N)) + (R(u_H^N) - R(u_h^N), (I - I_h^H)(u_h^N - v_h^N)).$$

From (6.1) we easily deduce

$$\begin{aligned} \mathcal{B}_h(u_h^N - u_H^N, u_h^N - v_h^N) &\lesssim J_h(\mathcal{M}) |u_h^N - v_h^N|_{1,h} + (J_h(\mathcal{M}) + J_H(\mathcal{M})) |u_h^N - v_h^N|_{1,h} \\ &\lesssim (J_h(\mathcal{M}) + J_H(\mathcal{M})) |u_h^N - v_h^N|_{1,h}. \end{aligned} \tag{6.5}$$

As far as the second term on the right-hand side in (6.4) is concerned, in view of the auxiliary problem (3.14) we can easily derive

$$||u_h^N - u_H^N||_{0,\Omega} \lesssim |u_h^N - u_H^N|_{1,h},$$

which gives rise to

$$\mathcal{B}_h(u_h^N - u_H^N, v_h^N - u_H^N) \lesssim ||\mathbf{A}^{\frac{1}{2}}\nabla_h(u_h^N - u_H^N)||_{0,\Omega}|u_H^N - v_h^N|_{1,h}. \tag{6.6}$$

Using (6.5),(6.6), applying the triangle inequality, Young's inequality and the inequality

$$J_h^2(\mathcal{M}) \le J_H^2(\mathcal{M}) + C_R||\mathbf{A}^{\frac{1}{2}}\nabla_h(u_h^N - u_H^N)||_{0,\Omega}^2,$$

that can be deduced as in Lemma 5.1, we get

$$\mathcal{B}_h(u_h^N - u_H^N, u_h^N - u_H^N) \lesssim J_H^2(\mathcal{M}) + \inf_{v_h^N \in V_h^N} |u_H^N - v_h^N|_{1,h}^2. \tag{6.7}$$

The assertion follows from (6.2) and (6.7). □

In view of the above arguments, we suggest **ANFEM II** through replacing the step MARK in **ANFEM I** by the following modification.

Modified MARK algorithm:
1. Given parameters $0 < \theta_1, \theta_2, \gamma < 1$, set $\mathcal{M}_0 = \emptyset$;
2. If $\gamma\eta_s^2(u_H^N) \ge J_H^2$, mark a set of edges $\mathcal{F} \subset \mathcal{E}_H$ with minimal cardinality such that

$$\theta_1\eta_s^2(u_H^N) \le \eta_s^2(u_H^N, \mathcal{F}).$$

Set $\mathcal{M} = \mathcal{M}_0 \cup \{\mathcal{F}\}$;
else mark a set of triangles $\hat{\mathcal{T}} \subset \mathcal{T}_H$ with minimal cardinality such that

$$\theta_2 J_H^2 \le J_H^2(\hat{\mathcal{T}}).$$

Set $\mathcal{M} = \mathcal{M}_0 \cup \{\hat{\mathcal{T}}\}$;
3. Mark further edges or triangles to avoid hanging nodes;
4. Output the marked edges or triangles $\mathcal{M}$:=MARK$(\mathcal{E}_H, \mathcal{T}_H)$.

The proof of the convergence result (5.4) for ANFEM II is fully similar to ANFEM I, so we omit it here. In order to get the optimality of ANFEM II, we need some notations from nonlinear approximation theory. Let $N_h = \dim(V_h^N)$, $\mathbb{H}_N = \{V_h^N | \dim(V_h^N) \le N\}$ and define the approximation class $\mathbb{A}_s$ as follows

$$\mathbb{A}_s := \{(u, f, D) | \quad |u, f, D|_s = \sup_{N \ge N_0} N^s \cdot \sigma(N; u, f, D) < +\infty\}, \tag{6.8}$$

where D denotes the data of the original problem and

$$\sigma(N; u, f, D) = \inf_{V_h^N \in \mathbb{H}_N} (|||u - u_h^N|||_h^2 + J_h^2).$$

THEOREM 6.1. *Assume $(u, f, D) \in \mathbb{A}_s$ and let $\{\mathcal{T}_k, V_k^N, u_k^N\}_{k\ge0}$ be a sequence of meshes, nonconforming finite element spaces and discrete solutions produced by* ANFEM II. *Moreover, set $\varepsilon_k := (|||u - u_k^N|||_{h_k}^2 + \beta J_k^2)^{\frac{1}{2}}$. Then, for sufficiently small initial mesh size h_0 the following optimal complexity estimate holds true*

$$\varepsilon_k \lesssim |u, f, D|_s (N_k - N_0)^{-s}. \tag{6.9}$$

Proof. For some $0 < \lambda < 1$ which will be specified later, we choose a triangulation $\mathcal{T}_{h^*}$ with minimal cardinality of $V_{h^*}^N$ such that

$$|||u - u_{h^*}^N|||_{h^*}^2 + \beta J_{h^*}^2 \leq \lambda(|||u - u_H^N|||_H^2 + \beta J_H^2).$$

By definition of $|u, f, D|_s$ we have

$$N_{h^*} \lesssim |u, f, D|_s^{\frac{1}{s}} (|||u - u_H^N|||_H^2 + \beta J_H^2)^{-\frac{1}{2s}}. \tag{6.10}$$

Suppose that $\mathcal{T}_{h^*}$ stems from a refinement of $\mathcal{T}_H$. In view of the properties of the step REFINE (cf. [6,30]), (6.10) can be replaced by

$$N_{h^*} - N_H \lesssim |u, f, D|_s^{\frac{1}{s}} (|||u - u_H^N|||_H^2 + \beta J_H^2)^{-\frac{1}{2s}}. \tag{6.11}$$

For the first case in the modified MARK algorithm, we prove that there exist $\mathcal{F}^* \subset \mathcal{E}_H$ and $0 < \theta < 1$ satisfying

$$\theta \eta_s^2(u_H^N) \leq \eta_s^2(u_H^N, \mathcal{F}^*). \tag{6.12}$$

In fact, using Lemma 4.1 we easily deduce

$$\begin{aligned}
|||u_{h^*}^N - u_H^N|||_{h^*}^2 = {} & |||u - u_H^N|||_H^2 - |||u - u_{h*}^N|||_{h^*}^2 - 2\mathcal{B}_{h^*}(u_{h^*}^N - u_H^N, u - u_{h*}^N) \\
& + 2(b \cdot \nabla_{h^*}(u_{h^*}^N - u_H^N), u - u_{h*}^N) \\
\geq {} & |||u - u_H^N|||_H^2 - |||u - u_{h*}^N|||_{h^*}^2 - 2C_0(J_{h^*} + J_H)(|||u - u_{h^*}^N|||_{h^*} \\
& + |||u - u_H^N|||_H) + 2(b \cdot \nabla_{h^*}(u_{h^*}^N - u_H^N), u - u_{h*}^N).
\end{aligned}$$

Then, by Young's inequality and (4.19), for some $\delta > 0$ we get

$$\begin{aligned}
& |||u_{h^*}^N - u_H^N|||_{h^*}^2 + C_1^{-1}(B^2 C_3^2 h_0^\alpha |||u - u_{h*}^N|||_{h*}^2 + h_0^\alpha |||u_{h^*}^N - u_H^N|||_{h*}^2) \\
\geq {} & |||u - u_H^N|||_H^2 - |||u - u_{h*}^N|||_{h^*}^2 - \frac{2C_0^2}{\delta}(2J_H^2 + C_R|||u_{h^*}^N - u_H^N|||_{h*}^2) \\
& - 2\delta(|||u - u_H^N|||_H^2 + |||u - u_{h*}^N|||_{h^*}^2),
\end{aligned} \tag{6.13}$$

where δ will be chosen later. The efficiency of the a posteriori error estimator (cf. Lemma 3.4) implies the existence of $C_7 > 0$ such that

$$\eta_s^2(u_H^N) \leq C_7 |||u - u_H^N|||_H^2.$$

From (6.13) we obtain

$$\begin{aligned}
& (1 + \frac{h_0^\alpha}{C_1} + \frac{2C_0^2 C_R}{\delta})|||u_{h^*}^N - u_H^N|||_{h^*}^2 \\
\geq {} & (1 - 2\delta)|||u - u_H^N|||_H^2 - (1 + 2\delta + \frac{B^2 C_3^2 h_0^\alpha}{C_1})|||u - u_{h^*}^N|||_{h^*}^2 - \frac{4C_0^2}{\delta} J_H^2 \\
\geq {} & (1 - 2\delta)|||u - u_H^N|||_H^2 - (1 + 2\delta + \frac{B^2 C_3^2 h_0^\alpha}{C_1})\lambda(|||u - u_H^N|||_H^2 + \beta J_H^2) - \frac{4C_0^2}{\delta} J_H^2 \\
\geq {} & (1 - 2\delta - (1 + 2\delta + \frac{B^2 C_3^2 h_0^\alpha}{C_1})\lambda) C_7^{-1} \eta_s^2(u_H^N) - ((1 + 2\delta + \frac{B^2 C_3^2 h_0^\alpha}{C_1})\lambda\beta + \frac{4C_0^2}{\delta}) J_H^2.
\end{aligned}$$

Using (4.19) and the quasi-discrete reliability (6.3) it follows that

$$\begin{aligned}
\eta_s^2(u_H^N, \mathcal{F}^*) \geq {} & \frac{1 - C_1^{-1} h_0^\alpha}{C_5} |||u_{h^*}^N - u_H^N|||_{h^*}^2 - J_H^2(\mathcal{M}^*) \\
& - \frac{B^2 C_3^2 h_0^\alpha}{2C_1 C_5}(|||u - u_{h^*}^N|||_{h^*}^2 + |||u - u_H^N|||_H^2).
\end{aligned} \tag{6.14}$$

Combining the above inequalities yields

$$\eta_s^2(u_H^N, \mathcal{F}^*) \geq \mu_1 \eta_s^2(u_H^N) - \mu_2 J_H^2 \geq (\mu_1 - \gamma\mu_2)\eta_s^2(u_H^N),$$

where

$$\mu_1 = \frac{(1-\frac{h_0^\alpha}{C_1})(1-2\delta-(1+2\delta+\frac{B^2C_3^2h_0^\alpha}{C_1})\lambda)}{C_5C_7(1+\frac{h_0^\alpha}{C_1}+\frac{2C_0^2C_R}{\delta})} - \frac{B^2C_3^2h_0^\alpha(1+\lambda)}{2C_1C_5C_7},$$

$$\mu_2 = \frac{(1+2\delta+\frac{B^2C_3^2h_0^\alpha}{C_1})\lambda\beta+\frac{4C_0^2}{\delta}}{1+\frac{h_0^\alpha}{C_1}+\frac{2C_0^2C_R}{\delta}} + \frac{B^2C_3^2h_0^\alpha\lambda\beta}{2C_1C_5}+1.$$

In order to obtain (6.12), we need to have $\mu_1 - \gamma\mu_2 \geq \theta$. A simple computation shows that this can be achieved by choosing δ, λ and γ according to

$$0 < \frac{\kappa_2 - \sqrt{\kappa_2^2 - 4\kappa_1\kappa_3}}{2\kappa_1} \leq \delta \leq \frac{\kappa_2 + \sqrt{\kappa_2^2 - 4\kappa_1\kappa_3}}{2\kappa_1} < \frac{1}{2},$$

where

$$\kappa_1 = 2(\gamma\lambda\beta + \frac{1+\lambda}{C_5C_7}) - O(h_0^\alpha)\ ,\ \kappa_2 = \frac{1-\lambda}{C_5C_7} - (1+\lambda\beta)\gamma - \theta - O(h_0^\alpha)\ ,\ \kappa_3 = 4\gamma C_0^2 + O(h_0^2).$$

There exist sufficiently small $\lambda_0 > 0$ and $\gamma_0 > 0$ such that $\kappa_2^2 - 4\kappa_1\kappa_3 \geq 0$ and $\delta < \frac{1}{2}$ for $0 < \lambda \leq \lambda_0$ and $0 < \gamma \leq \gamma_0$. We set $N_k := N_H$. Since $\mathcal{F}$ has been chosen with minimal cardinality to satisfy (5.1), we arrive at

$$N_{k+1} - N_k \lesssim \#\mathcal{F} \lesssim \#\mathcal{F}^* \lesssim N_{h^*} - N_k \lesssim |u, f, D|_s^{\frac{1}{s}} \varepsilon_k^{-\frac{1}{s}}.$$

For the second case in the modified MARK algorithm, we prove that there exist $\mathcal{M}^* \subset \mathcal{T}_H$ and $0 < \theta' < 1$ satisfying

$$\theta' J_H^2 \leq J_H^2(\mathcal{M}^*). \tag{6.15}$$

In fact, observing

$$J_{h^*}^2 \geq \gamma_0 J_H^2 - C_R |||u_{h^*}^N - u_H^N|||_{h^*}^2,$$

where γ_0 and C_R are the same as in Lemma 5.1, the global reliability

$$|||u - u_H^N|||_H^2 \leq C_8(\eta_s^2(u_H^N) + J_H^2),$$

and the quasi-discrete reliability (6.3), we get

$$\begin{aligned}\beta J_H^2 &\geq \lambda^{-1}(|||u - u_{h^*}^N|||_{h^*}^2 + \beta J_{h^*}^2) - |||u - u_H^N|||_H^2 \\ &\geq \lambda^{-1}|||u - u_{h^*}^N|||_{h^*}^2 + \lambda^{-1}\beta(\gamma_0 J_H^2 - C_R|||u_{h^*}^N - u_H^N|||_{h^*}^2) - C_8(\eta_s^2(u_H^N) + J_H^2).\end{aligned}$$

In view of (6.14) it follows that

$$(\gamma_0 + \frac{C_RC_5}{1-\frac{h_0^\alpha}{C_1}})J_H^2(\mathcal{M}^*) \geq \vartheta_1 J_H^2 + \vartheta_2 |||u - u_{h^*}^N|||_{h^*}^2,$$

where

$$\vartheta_1 = \gamma_0 - \lambda - \frac{C_R C_5}{(1-\frac{h_0^\alpha}{C_1})\gamma} - C_8(1+\frac{1}{\gamma})(\frac{\lambda}{\beta} + \frac{C_R B^2 C_3^2 h_0^\alpha}{2C_1(1-\frac{h_0^\alpha}{C_1})}),$$
$$\vartheta_2 = \frac{1}{\beta} - O(h_0^{\alpha+2}).$$

Obviously, $\vartheta_2 > 0$ for h_0 sufficiently small. Hence, (6.15) is guaranteed, if

$$\vartheta_1 \geq \theta'(\gamma_0 + \frac{C_R C_5}{1-\frac{h_0^\alpha}{C_1}}),$$

which holds true by specifying λ according to

$$0 < \lambda \leq \frac{(1-\frac{h_0^\alpha}{C_1})\gamma_0(1-\theta')C_1 - (\theta'+\frac{1}{\gamma})C_R C_5 C_1 - \frac{1}{2}(1+\frac{1}{\gamma})h_0^\alpha B^2 C_8 C_3^2}{(1+\frac{C_8}{\beta}(1+\frac{1}{\gamma}))(C_1 - h_0^\alpha)}.$$

Since $\mathcal{M}$ has been chosen with minimal cardinality to satisfy (5.2), we obtain

$$N_{k+1} - N_k \lesssim \#\mathcal{M} \lesssim \#\mathcal{M}^* \lesssim N_{h^*} - N_k \lesssim |u,f,D|_s^{\frac{1}{s}} \varepsilon_k^{-\frac{1}{s}}.$$

Therefore, the modified MARK algorithm always yields

$$N_k - N_0 = \sum_{l=0}^{k-1}(N_{l+1} - N_l) \lesssim \sum_{l=0}^{k-1} \varepsilon_l^{-\frac{1}{s}} \cdot |u,f,D|_s^{\frac{1}{s}}. \tag{6.16}$$

In view of the convergence result $\varepsilon_{l+1} \leq \rho^{\frac{1}{2}} \varepsilon_l$, we have

$$\varepsilon_l^{-\frac{1}{s}} \leq \varepsilon_{k-1}^{-\frac{1}{s}} \rho^{\frac{k-1-l}{2s}},$$

and (6.16) implies

$$N_k - N_0 \lesssim |u,f,D|_s^{\frac{1}{s}} (\sum_{l=0}^{k-1} \rho^{\frac{k-1-l}{2s}}) \cdot \varepsilon_{k-1}^{-\frac{1}{s}} \lesssim |u,f,D|_s^{\frac{1}{s}} \varepsilon_k^{-\frac{1}{s}},$$

which results in (6.9). □

7. Numerical results. In this section, we present a detailed documentation of numerical results based on the application of ANFEM I and ANFEM II to show how the meshes are generated adaptively and how the estimators and the errors behave due to effects from different coefficients in (1.2), the lack regularity of solutions and different domains. The implementation of ANFEM I and II is based on the FFW toolbox as described in [14].

For ease of notation, we refer to $(\eta_s + J_h)_{\mathrm{k}}$ as the sum of the a posteriori error estimators computed in step k of the ANFEM algorithms, $\mathrm{DOF_k}$ stands for the number of degrees of freedom in $\mathcal{T}_k$, and $\mathrm{M}_E, \mathrm{M}_T$ denote the total number of marked edges and marked triangles in each iteration, respectively. In the following experiments, the parameters in the ANFEM algorithms have been chosen according to $\theta_1 = 0.5$ and $\theta_2 = 0.5$. The figures displaying the convergence history are all plotted in log-log coordinates.

Example 7.1. Consider the PDE (1.2) with Dirichlet boundary conditions on the L-shaped domain $\Omega = (-1,0)\times(-1,1)\cup[0,1]\times(0,1)$, where the the exact solution is given in polar coordinates by

$$u(r,\theta) = r^{\frac{2}{3}} \sin(\frac{2}{3}\theta).$$

Example 7.1.1. In the first experiment, the coefficients in (1.2) are chosen as follows

$$\mathbf{A} = \epsilon I \quad , \quad \mathbf{b} = \binom{r\cos\theta}{r\sin\theta} \ , \qquad c = r^{\frac{1}{4}},$$

where $\epsilon = 10^{-2}$ and I stands for the identity matrix.

We note that the problem is singularly perturbed which will be reflected by the convergence behavior of the algorithms at the beginning of the adaptive process.

Table 7.1
Example 7.1.1: ANFEM I (left) and ANFEM II (right)

k	$\mathrm{DOF_k}$	$(\eta_s + J_h)_\mathrm{k}$	M_E	M_T	k	$\mathrm{DOF_k}$	$(\eta_s + J_h)_\mathrm{k}$	M_E	M_T
18	13902	0.041	1782	447	21	25443	0.028	3903	0
19	23884	0.030	3081	774	22	41620	0.021	6329	0
20	40585	0.023	5464	1634	23	66527	0.017	11207	0
21	72239	0.017	10084	2904	24	112127	0.013	18119	0
22	130184	0.013	17544	5140	25	184672	0.010	29525	0
23	226232	0.00962	32144	9490	26	300700	0.0077	51868	0

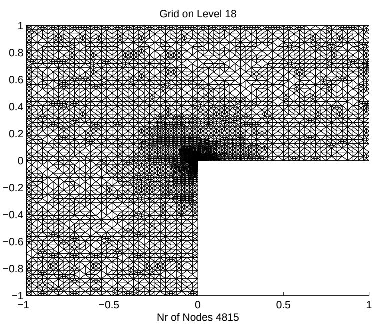

FIG 7.1: Mesh on level 18 of ANFEM I (Example 7.1.1)

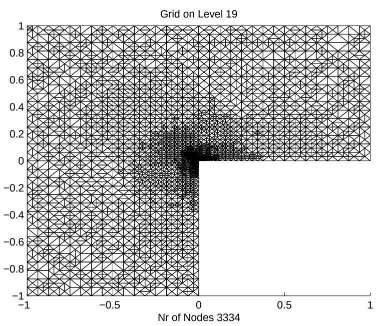

FIG 7.2: Mesh on level 19 of ANFEM II (Example 7.1.1)

Table 7.1 contains the results beginning with step 18 of ANFEM I and step 21 of ANFEM II ($\gamma = 0.6$ in ANFEM II). Figures 7.1 and 7.2 display the meshes after 18 iterations of ANFEM I and 19 iterations of ANFEM II. Figure 7.3 shows the energy error norm of ANFEM II (cf. (2.7)) and the a posteriori error estimator as functions of the number of degrees of freedom. The results support the optimal convergence $\propto \mathrm{DOF}^{-\frac{1}{2}}$ for ANFEM II.

Example 7.1.2. In the second experiment, we only change the coefficient $\mathbf{A}$ to $\mathbf{A} = \mathrm{I}$, the other coefficients and parameters remain the same. We display the mesh after

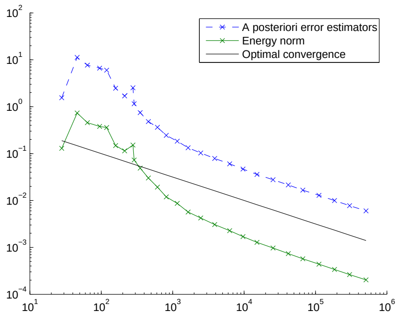

FIG 7.3: Convergence history of ANFEM II (Example 7.1.1)

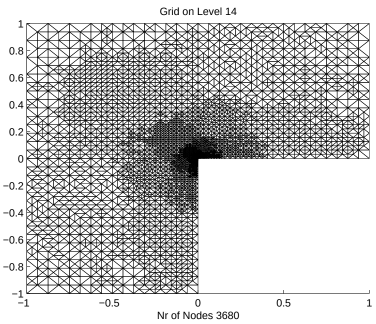

FIG 7.4: Mesh on level 14 of ANFEM II (Example 7.1.2)

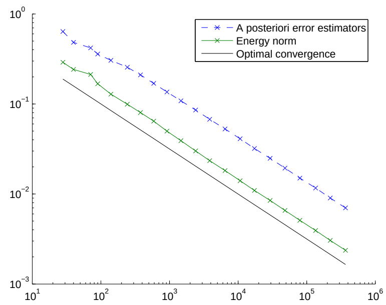

FIG 7.5: Convergence history of ANFEM II (Example 7.1.2)

14 iterations of ANFEM II and the corresponding convergence history in Figures 7.4 and 7.5.

Compared to Example 7.1.1, here the elliptic problem is not singularly perturbed which is reflected by the steady decrease of the error and the estimator from the very beginning of the adaptive process.

EXAMPLE 7.2. We consider a convection-dominated convection-diffusion problem with Dirichlet boundary conditions on the square $\Omega = (0,1)^2$ (cf. experiment 2 in [28]), where the coefficients in (1.2) are chosen according to

$$\mathbf{A} = \epsilon I, \epsilon = 10^{-3}; \qquad \mathbf{b} = \begin{pmatrix} y \\ 0.6-x \end{pmatrix}, \qquad c = f = 0,$$

and the Dirichlet data g on $\partial\Omega$ is given by

$$g(x,y) = \begin{cases} 1, & \{0.3+\tau \le x \le 0.6-\tau, y=0\}\ , \\ 0, & \partial\Omega \backslash \{0.3 \le x \le 0.6, y=0\}\ , \\ \text{linear}, & \{0.3 \le x \le 0.3+\tau, y=0\} \text{ or } \{0.6-\tau \le x \le 0.6, y=0\}\ . \end{cases}$$

Here, the parameter τ can be chosen freely in the experiment. The following results are based on $\tau = 0.003$.

For $\gamma = 0.5$ in ANFEM II, Table 7.2 contains the results beginning with step 16 of ANFEM I and step 23 of ANFEM II. Figures 7.6 and 7.7 display the mesh after 25

Table 7.2
Example 7.2: ANFEM I (left) and ANFEM II (right)

k	$\mathrm{DOF_k}$	$(\eta_s + J_h)_\mathrm{k}$	M_E	M_T	k	$\mathrm{DOF_k}$	$(\eta_s + J_h)_\mathrm{k}$	M_E	M_T
16	11082	1.10	374	1210	23	15867	0.74	2382	0
17	20750	0.79	1084	2484	24	26273	0.57	4177	0
18	39217	0.55	3267	4453	25	44731	0.44	6710	0
19	73451	0.38	7738	9230	26	73513	0.34	11907	0
20	141645	0.27	16773	16179	27	124657	0.26	19880	0
21	260916	0.20	32448	31890	28	211231	0.20	32603	0

iterations of ANFEM II and the convergence history. The experimental results also support the optimal convergence $\propto \mathrm{DOF}^{-\frac{1}{2}}$ of ANFEM II.

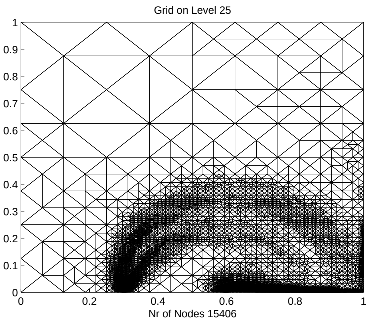

FIG 7.6: Mesh on level 25 of ANFEM II (Example 7.2)

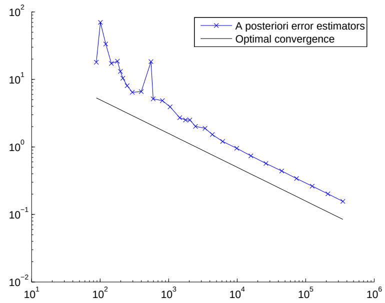

FIG 7.7: Convergence history of ANFEM II (Example 7.2)

Remark: Experiments 7.1.1, 7.1.2 and 7.2 all start from a uniform mesh coarser than required in theory. Though the errors increase at the beginning of the iteration (see experiments 7.1.1 and 7.2), they decrease after an initial phase and asymptotically result in an optimal convergence rate.

8. Extensions. For model problems with discontinuous coefficients, Bernardi and Verfürth [6] derived a new a posteriori error estimator for AFEM based on conforming finite element method, which was thereafter improved by Chen and Dai [17]. For nonconforming schemes, a robust a posteriori error estimator with respect to jumps in coefficients was studied by Ainsworth in [1]. In the sequel, we extend our results to the nonsymmetric and indefinite problem (1.2),(1.3) by allowing $\mathbf{A}$ to have large jumps in Ω.

We require the initial mesh to be chosen such that $\mathbf{A}$ is piecewise constant on $\mathcal{T}_0$, and we also make the following assumption instead of (1.4): for all $T \in \mathcal{T}_h$ there exist constants λ_T and Λ_T satisfying

$$\lambda_T ||\xi||_{0,T}^2 \leq (\mathbf{A}\xi, \xi)_T \leq \wedge_T ||\xi||_{0,T}^2 \quad , \quad \xi \in L_2(T)^2, \tag{8.1}$$

with the ratio $\gamma_T = \Lambda_T / \lambda_T$ being uniformly bounded with respect to the family of meshes.

We define the new a posteriori error estimators as follows: the new edge residual of the tangential component is given by

$$\widetilde{\eta}_{s,E}(u_h^N) := \begin{cases} \Lambda_E^{\frac{1}{2}} h_E^{\frac{1}{2}} ||[\frac{\partial u_h^N}{\partial s}]||_{0,E} \ , \ E \in \mathcal{E}_h^0 \\ \Lambda_E^{1/2} \| \frac{\partial u_h^N}{\partial s} \|_{0,E} \ , \ E \in \mathcal{E}_h^{\partial\Omega} \end{cases} , \tag{8.2}$$

where $\Lambda_E := \max(\Lambda_T, \Lambda_{T'}), T \cap T' = E$, and $\Lambda_E = \Lambda_T, T \cap \partial\Omega = E$. We further set

$$\widetilde{\eta}_s^2(v_h^N, \mathcal{F}) := \sum_{E \in \mathcal{F}} \widetilde{\eta}_{s,E}^2(u_h^N) \quad , \quad \mathcal{F} \subseteq \mathcal{E}_h. \tag{8.3}$$

The new edge residual of the normal component is given by

$$\widetilde{\eta}_{\nu,E}(u_h^N) := \lambda_E^{-\frac{1}{2}} h_E^{\frac{1}{2}} ||[\mathbf{A}\nabla u_h^N] \cdot \nu_E||_{0,E} \quad , \quad E \in \mathcal{E}_h^0, \tag{8.4}$$

where $\lambda_E = \min\{\lambda_T, \lambda_{T'}\}, \partial T \cap \partial T' = E$. We set

$$\widetilde{\eta}_\nu^2(u_h^N, \mathcal{F}) := \sum_{E \in \mathcal{F}} \widetilde{\eta}_{\nu,E}^2(u_h^N) \quad , \quad \mathcal{F} \subseteq \mathcal{E}_h^0. \tag{8.5}$$

Finally, the new volume and oscillation terms are defined according to

$$\widetilde{J}_T(u_h^N) := \lambda_T^{-\frac{1}{2}} h_T ||R_T(u_h^N)||_{0,T} \quad , \quad T \in \mathcal{T}_h, \tag{8.6}$$

$$\widetilde{\text{osc}}_T(u_h^N) := \lambda_T^{-\frac{1}{2}} h_T ||R_T(v_h^N) - \overline{R_T(u_h^N)}||_{0,T} \quad , \quad T \in \mathcal{T}_h, \tag{8.7}$$

and we set

$$\widetilde{J}_h^2(u_h^N, \mathcal{M}) := \sum_{T \in \mathcal{M}} \widetilde{J}_T^2 \quad , \quad \mathcal{M} \subseteq \mathcal{T}_h, \tag{8.8}$$

$$\widetilde{\text{osc}}_h^2(v_h^N, \mathcal{M}) := \sum_{T \in \mathcal{M}} \widetilde{\text{osc}}_T^2(v_h^N) \quad , \quad \mathcal{M} \subseteq \mathcal{T}_h. \tag{8.9}$$

For the new a posteriori error estimators we use the same abbreviations as before. Following the lines of proof of the reliability and the efficiency of the old estimators, we can derive the same results except that the efficiency for $\widetilde{\eta}_{s,E}(u_h^N)$ and $\widetilde{J}_T(u_h^N)$ has to be modified as follows.

LEMMA 8.1. *Let $\widetilde{\eta}_{s,E}(u_h^N)$ and $\widetilde{J}_T(u_h^N)$ be given by* (8.2) *and* (8.6). *Then, there holds*

$$\widetilde{\eta}_{s,E}(u_h^N) \lesssim \Lambda_E^{\frac{1}{2}} ||\nabla_h(u - u_h^N)||_{0,w_E}, \tag{8.10}$$

$$\widetilde{J}_T(u_h^N) \lesssim \lambda_T^{-\frac{1}{2}} ||u - u_h^N||_{1,T} + \widetilde{\text{osc}}_h(T). \tag{8.11}$$

The discrete local efficiency reads as follows

$$\widetilde{\eta}_{s,E}(u_h^N) \lesssim \Lambda_E^{\frac{1}{2}} ||\nabla_h(u_h^N - u_H^N)||_{0,w_E}. \tag{8.12}$$

We have the same quasi-orthogonality and volume term reduction as in Theorem 4.1 and Lemma 5.1. However, for the convergence results with respect to ANFEM I and ANFEM II, the error reduction rate $\widetilde{\rho}$ is different from the rate ρ in Theorem 5.1. Based on the discrete local efficiency (8.12) and the proof of Theorem 5.1, we can show

$$0 < \widetilde{\rho} = \max\{\Lambda_0 - \frac{\beta_0 \delta_0}{\Lambda}, \gamma_1\} < 1,$$

where

$$\Lambda = \max_{E\in\mathcal{E}_h} \Lambda_E \quad , \quad 0 < \rho_1 < \gamma_1 < 1 \quad , \quad 0 < \wedge_0 - \frac{\beta_0\delta_0}{\Lambda} < 1,$$

provided there holds

$$O(h_0^\alpha) < \frac{\beta_0\delta_0}{2\Lambda + \beta_0\delta_0}.$$

Following Lemma 6.1 and Theorem 6.1, we can also derive the optimal complexity for ANFEM II.

Observing the appearance of Λ in the definition of $\widetilde{\rho}$, we see that the error reduction rate $\widetilde{\rho}$ will increase, if the jumps in the coefficient $\mathbf{A}$ get large. This behavior can be observed in the following experiment as well.

EXAMPLE 8.1. We choose the example from [18, 29]. Let $\Omega = (-1,1)^2$, and choose

$$\begin{aligned}
&\mathbf{A} = \epsilon R I \text{ in the first and third quadrants,}\\
&\mathbf{A} = \epsilon I \text{ in the second and fourth quadrants,}\\
&\mathbf{b} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \quad , \quad c = 0, \quad \text{where } \epsilon = 10^{-2} \ , \ R = const.
\end{aligned}$$

All other data are such that the exact weak solution of (1.2) is given in polar coordinates by $u(r,\theta) = r^\tau \mu(\theta)$, where

$$\mu(\theta) = \begin{cases} \cos((\frac{\pi}{2}-\sigma)\tau)\cdot \cos((\theta-\frac{\pi}{2}+\rho)\tau), & \text{if } 0 \le \theta \le \frac{\pi}{2} \ ,\\ \cos(\rho\tau)\cdot \cos((\theta-\pi+\sigma)\tau), & \text{if } \frac{\pi}{2} \le \theta \le \pi \ ,\\ \cos(\sigma\tau)\cdot \cos((\theta-\pi-\rho)\tau), & \text{if } \pi \le \theta \le \frac{3}{2}\pi \ ,\\ \cos((\frac{\pi}{2}-\rho)\tau)\cdot \cos((\theta-\frac{3\pi}{2}-\sigma)\tau), & \text{if } \frac{3}{2}\pi \le \theta \le 2\pi \ . \end{cases}$$

Here, the numbers τ, ρ, σ, R satisfy the nonlinear equations

$$\begin{cases} \text{R=-}\tan((\frac{\pi}{2}-\sigma)\tau)\cdot\cot(\rho\tau),\\ 1/R\text{=-}\tan(\rho\tau)\cdot\cot(\sigma\tau),\\ \text{R=-}\tan(\sigma\tau)\cdot\cot((\frac{\pi}{2}-\rho)\tau),\\ 0<\tau<2,\\ \max\{0,\pi\tau-\pi\} < 2\tau\rho < \min\{\pi\tau,\pi\},\\ \max\{0,\pi-\pi\tau\} < -2\tau\sigma < \min\{\pi\tau, 2\pi-\pi\tau\}. \end{cases}$$

The solution u is in H^{1+s} with $s < \tau$. For $\tau = 0.1$, the above nonlinear equations are solved by Newton's method, and we obatin

$$R \approx 161.4476387975881 \quad , \quad \rho = \frac{\pi}{4} \quad , \quad \sigma \approx -14.92256510455152.$$

In the following experiments, we choose the parameters in ANFEM I and ANFEM II as $\theta_1 = 0.6$, $\theta_2 = 0.2$, and $\gamma = 0.6$ in ANFEM II. We first solve this problem by ANFEM I using the new a posteriori error estimators. Figures 8.1 and 8.2 display the mesh after 33 iterations and the convergence history, respectively. Figure 8.1 shows that the mesh generated by ANFEM I is not optimal, since the mesh is not always refined around the singular point. Intrinsically, the volume term does not always reflect the singularity. The convergence history also indicates that the complexity of ANFEM I is not optimal.

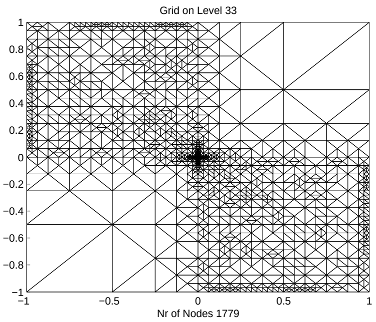

FIG 8.1: Mesh on level 33 of ANFEM I by the new a posteriori error estimator

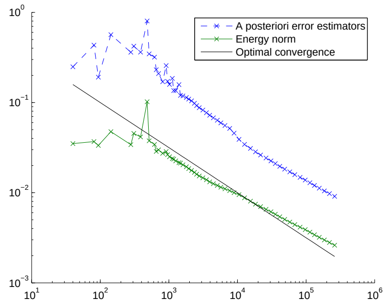

FIG 8.2: Convergence history of ANFEM I by the new a posteriori error estimator

As far as the performance of ANFEM II is concerned, Figure 8.3 is based on 50174 DOFs generated by the old a posteriori error estimator, whereas Figure 8.4 is based on 50942 DOFs generated by the new estimator. The energy error of the finite element solution for the mesh in Figure 8.4 is much less than the one for the mesh in Figure 8.3. It clearly shows that the new a posteriori error estimator provides a better approximation of the true solution than the old one. Figure 8.5 displays the convergence history of ANFEM II by the new a posteriori error estimator indicating optimal complexity.

TABLE 8.1
Example 8.1

τ	k	$\mathrm{DOF}_{\mathrm{k}}$	CPU time	$(\eta_s + J_h)_{\mathrm{k}}$
0.02	175	43805	326.03s	0.057
0.1	54	5870	11.75s	0.053
0.3	40	2348	6.06s	0.049

Finally, we show that the error reduction rate $\widetilde{\rho}$ increases when the jumps in the coefficient $\mathbf{A}$ get large, i.e., when the singularity of the solution becomes stronger. We will compare three situations by choosing $\tau = 0.02$, $\tau = 0.1$, and $\tau = 0.3$. For $\tau = 0.02$, the parameters R, ρ, σ are given by

$$R \approx 4052.1806954768103 \quad , \quad \rho = \frac{\pi}{4} \quad , \quad \sigma \approx -77.754418176347386,$$

whereas for $\tau = 0.3$ the parameters R, ρ, σ are as follows

$$R \approx 17.3497221747152608 \quad , \quad \rho = \frac{\pi}{4} \quad , \quad \sigma \approx -4.4505895925855404.$$

Table 8.1 shows that the iteration number increases for increasing jumps in the coefficient $\mathbf{A}$ which confirms our theoretical result.

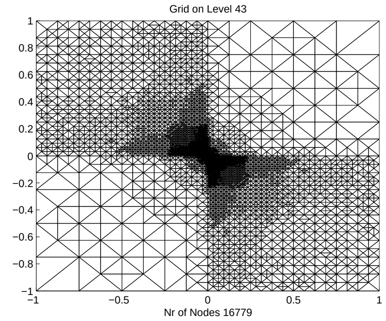

FIG 8.3: Mesh on level 43 of ANFEM II by the old a posteriori error estimator

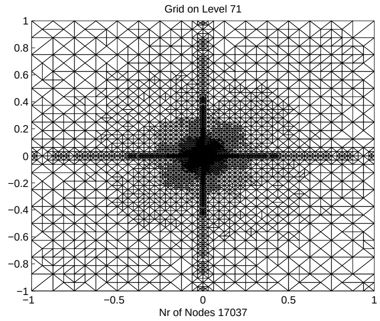

FIG 8.4: Mesh on level 71 of ANFEM II by the new a posteriori error estimator

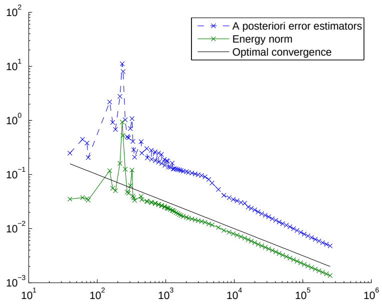

FIG 8.5: Convergence history of ANFEM II by the new a posteriori error estimator

REFERENCES

[1] M. Ainsworth, Robust a posteriori error estimation for nonconforming finite element approximation, SIAM J. Numer. Anal., **42**(2005), 2320-2341.

[2] M. Ainsworth and J.T. Oden, A posteriori error estimation in finite element analysis, John Wiley & Sons, Inc., New York, 2000.

[3] T. Arbogsat and Z.X. Chen, On the implementation of mixed methods as nonconforming methods for second-order elliptic problems, Math. Comp., **64**(1995), 943-972.

[4] C. Bahriawati and C. Carstensen, Three Matlab implementations of the lowest-order Raviart-Thomas MFEM with a posteriori error control, Computational Methods in Applied Mathematics, **5**(2005), No.4, 333-361.

[5] R. Becker, S. Mao, and Z. Shi, A convergent adaptive nonconforming finite element method with optimal complexity, in preparation.

[6] C. Bernardi and R. Verfürth, Adaptive finite element methods for elliptic equations with nonsmooth coefficients, Numer. Math., **85**(2000), 579-608.

[7] P. Binev, W. Dahmen, and R. DeVore, Adaptive finite element methods with convergence rates, Numer. Math., **97**(2004), 219-268.

[8] C. Carstensen, Quasi-interpolation and a posteriori error analysis in finite elements methods, RAIRO Model. Math. Anal. Numer., **33**(1999), 1187-1202.

[9] C. Carstensen and R.H.W. Hoppe, Convergence analysis of an adaptive nonconforming finite element emthod, Numer. Math., **103**(2006), 251-266.

[10] C. Carstensen and R.H.W. Hoppe, Error reduction and convergence for an adaptive mixed finite element method, Math. Comp., **75**(2006), 1033-1042.

[11] C. Carstensen, J. Hu, and A. Orlando, Framework for the a posteriori error analysis

of nonconforming finite elements, SIAM J. Numer. Anal., **45**(2007), 68-82.
[12] C. Carstensen and J. Hu, A unifying theory of a posteriori error control for nonconforming finite element methods, Numer. Math, **107**(2007), 473-502.
[13] C. Carstensen, S. Bartels, and S. Jansche, A posteriori error estimates for nonconforming finite element methods, Numer. Math, **92**(2002), 233-256.
[14] C. Carstensen, J. Gedicke, R. Huth, J. Reininghaus, D. Guenther, H. Mellmann, and S. Wiedemann, FFW Documentation, Humboldt University of Berlin, Germany, 2007.
[15] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Sibert, Quasi-optimal convergence rate for an adaptive finite element method, to appear, 2007.
[16] J. Chen and L. Li, Convergence and domain decomposition algorithm for nonconforming and mixed methods for nonselfadjoint and indefinite problems, Comput. Methods Appl. Mech. Engrg., **173**(1999), 1-20.
[17] L. Chen, M. Holst, and J. Xu, Convergence and optimality of adaptive mixed finite element methods, to appear in Math. Comp.
[18] Z. Chen and S. Dai, On the efficiency of adaptive finite element methods for elliptic problems with discontinous coefficients, SIAM J. Sci. Comput., **24**(2002), 443-462.
[19] Z.X. Chen, D.Y. Kwak, and Y.J. Yon, Multigrid algorithms for nonconforming and mixed methods for nonsymmetric and indefinite problems, SIAM J. Sci. Comput., **19**(1998), 502-515.
[20] P.G. Ciarlet, The finite element method for elliptic problems, North-Holland, Amsterdam, 1978.
[21] E. Dari, R.G. Duran, C. Padra, and V. Vampa, A posteriori error estimators for nonconforming finite element methods, RAIRO Model. Math. Anal. Numer., **30**(1996), 385-400.
[22] J. Douglas, Jr. and J.E. Robert, Global estimates for mixed methods for second order elliptic equations, **44**(1985), 39-51.
[23] D. Gilbarg and N.S. Trudinger, Elliptic partial differential equations of second order, Springer, Berlin-Heidelberg-New York, 1998.
[24] W. Hackbusch, Elliptic differential equations theory and numerical treatment, Springer, Berlin-Heidelberg-New York, 2003.
[25] R.H.W. Hopppe and B. Wohlmuth, Element-oriented and edge-oriented local error estimators for nonconforming finite element methods, RAIRO Model. Math. Anal. Numer., **30**(1996), 237-263.
[26] J. Hu and J. Xu, Convergence of adaptive conforming and nonconforming finite element methods for the perturbed Stokes equation, in preparation.
[27] F. Jochmann, An H^s regularity result for the gradient of solutions to elliptic equations with mixed boundary conditions, J. Math. Anal. Appl., **238**(1999), 429-450.
[28] K. Mekchay and R.H. Nochetto, Convergence of adaptive finite element methods for general second order linear elliptic PDEs, SIAM J. Numer. Anal., **43**(2005), 1803-1827.
[29] P. Morin, R.H. Nochetto and K.G. Siebert, Convergence of adaptive finite element methods, SIAM Review, **44**(2002), 631-658.
[30] P. Oswald, Integrid transfer operators and multilevel preconditioners for nonconforming discretizations, Appl. Numer. Math., **23**(1996), 139-158.
[31] A.H. Schatz, An observation concerning Ritz-Galerkin methods with indefinite bilinear forms, Math. Comp., **28**(1974), 959-962.
[32] F. Schieweck, A posteriori error estimates with post-processing for nonconforming finite elements, ESAIM Math. Mod. Numer. Anal., **36**(2002), 489-503.
[33] R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, Math. Comp., **54**(1990), 483-493.
[34] Z. Shi and X. Xu, On a note for nonconforming multigrid for nonsymmetric and indefinite problems, Chinese J. Numer. Math. Appl., **22**(2000), 102-108.
[35] R. Stevenson, Optimality of a standard adaptive finite element method, Foundations

of Computional Mathematics, **2**(2007), 245-269.

[36] R. Verfürth, A review of a posteriori error estimation and adaptive mesh-refinement technique, Wiley-Teubner, Chichester, 1996.

[37] S.P. Wong, Preconditioning nonconforming finite element methods for treating Dirichlet boundary conditions, II, Numer. Math., **62**(1992), 413-437.

[38] S.P. Wong, Preconditioning P1 nonconforming finite elements: condition numbers and singular value distributions, Numer. Math., **65**(1993), 253-271.

[39] Y.J. Yon and D.Y. Kwak, Nonconforming multigrid method for nonsymmetric and indefinite problems, Comput. Math. Appl., **30**(1995), 1-7.

[40] Y.J. Yon and D.Y. Kwak, Two-level additive Schwarz preconditioners for P1 nonconforming finite elements for nonsymmetric and indefinite problems, Appl. Math. Comput., **87**(1997), 1-14.